Advanced Pre-5G Capacity Enablers

Hooshmand, Ali

Awarding institution:
King's College London

Advanced Pre-5G Capacity Enablers

by

Ali Hooshmand

A thesis submitted to King's College London

in fulfilment of the requirements for the degree of

Doctor of Philosophy

March, 2016

Acknowledgements

I would like to express my special gratitude and appreciation to my supervisor Professor Hamid Aghvami who supported me during my entire research. His priceless guidance and encouragements motivated me throughout.

I would also like to thank the academic staff at the Department of Informatics formerly known as the Centre for Telecommunication Research (CTR). Their valuable support and motivational assistance during my Ph.D. years contributed greatly to the fulfilment of my research. I specially like to thank Prof. Dohler for his continued support and advice.

I also acknowledge all my friends and colleagues in the college specially Dr. Hadi Saki for the fruitful and excellent meetings and comments from which I learned a great deal. My special appreciation and thanks go to my family. I am genuinely grateful for them to motivate and support me during the challenges and hardships I faced. Their exceptional patience and love helped me in striving towards this academic goal of mine.

Abstract

Mobile user subscriptions have grown at an exponential rate recently. This trend is projected to continue at an even faster pace. Smart devices with an immense variety of applications are at the core of such drive. This has lead to the user expectations for receiving higher quality mobile broadband services to rise rapidly. With the current mobile communication technologies, satisfying an increasing demand for unprecedented volumes of data rate for users is impractical. The impairing effects of inter-cell interference (ICI) is identified as one of the most important hurdles in the way of the realisation of this gaol. Evolution of the current mobile networks into future versions that can adequately cope with the growing requirements entails utilisation of more spectrally efficient technologies. Envisioned for fulfilling high user Quality of Experience (QoE) in the future generation of mobile networks, Long Term Evolution (LTE) Advanced was introduced by the Third Generation Partnership Project (3GPP) as the candidate technology for 4G. This involves the refinement of some of the existing key technologies available in LTE and also the introduction of new innovative solutions. Next, the contributions and novelties of this thesis will be outlined.

Major Considerations and Contributions

Effort has been made to adhere to realistic scenario settings and assumptions in proposing guidelines and algorithms throughout the above research phases. The contribution of this Ph.D. dissertation is on the study of a number of can-

didate capacity-enabling technologies prior to 5G. There have been two major approaches to this. First, enhancing capacity is investigated from a Radio Resource Management (RRM) perspective for the DL of a Coordinated MultiPoint (CoMP)-enabled system. Carrier aggregation dynamics are then looked at in resource allocation for a single transmission multi-carrier network. Second, methods and strategies are investigated and proposed in cell association under CoMP to achieve maximum system capacity in a Heterogeneous Network (HetNet). Major considerations in achieving maximum system capacity through resource allocation (2 technical chapters) were pursued by:

- Providing minimum user Quality of Service (QoS);
- Enabling channel-aware adaptive power allocation (in case of a multi-carrier system);
- Allowing for adaptive modulation and coding (AMC) for link adaptation (for CoMP systems);
- Employing priority time and frequency scheduler;
- Taking into account unique propagation characteristics for transmitting at different carrier frequencies (in case of a multi-carrier system);
- Considering a component carrier (CC) load balancing platform for packet scheduling.

Furthermore, maximum network capacity realisation studied in the last research process in the context of HetNet cell association reflects on:

- Decoupled access for a cooperating and heterogeneous complex of evolved Node Bs (eNBs);
- Considering fronthaul rate limitations in the UL/DL;

- Cell association in the light of a controlled imbalance in UL/DL rate.

This Ph.D. thesis focuses its contributions to researching into below technologies and solutions intended for the forthcoming wireless communication generations through analysis and simulations through MATLAB.

CoMP as a competent and exciting technology offers a solution to manage the adverse impact of ICI on the cell-edge users' performance. Simultaneous transmission on the same time-frequency resources is facilitated through base station cooperation as a result of which, higher cell-edge signal-to-interference-and-noise ratio (SINR) levels are experienced.
However, even though near Shannon-limit capacity and spectral efficiencies have already been obtained through many available solutions, the demand for future higher data rates are still far from being achieved. To this end, acquisition of more bandwidth remains a uncompromising necessity.

Only supported in LTE-Advanced (LTE-A), **carrier aggregation (CA)** is another capable technology through which simultaneous transmission of data is made possible on multiple carrier frequencies. Bandwidth fragments are combined in transmission that may belong to different carriers and this brings flexibility due to unavailability of large spectrum.

On the other hand, *ultra dense* cell deployments have recently been an attractive subject for research in the community. This is because their ability to significantly enhance system capacity has just been unveiled. Contrary to the conventional arrangements, **Heterogeneous Networks (HetNets)** utilise multiple tiers of base stations distinguished by their transmit power levels. Small cells are introduced to help increase hotspot connectivity whilst providing coverage for macro base station blind spots. Very attractive deployment scenarios arise from HetNets and whilst the economical feasibility for deployment and management costs is another supplementary reason to their appeal.

By the means of general and critical literature review for each stage, the above topics constitute the novelties and contributions of the research phases presented in this dissertation. The unifying objective for all the stages of research in this work is the *realisation of maximum system capacity*. Strategies and algorithms for resource allocation in the context of CoMP and CA were examined and researched in the 1^{st} and 2^{nd} technical chapters followed by guidelines and methods studied in the 3^{rd} chapter from a cell association standpoint. Gaps in the existing literature were identified and effort was made to present more relevant, realistic and sufficient design strategies for resource allocation. This has been pursued whilst addressing a more comprehensive list of design challenges recognised in the state-of-the-art phase.

Contents

List of Figures

List of Tables

Acronyms

3GPP Third Generation Partnership Project. 3, 18, 19, 60

ABS Almost Blank Subframe. 54

AMC adaptive modulation and coding. 20, 64, 68, 76

ARPs allocation and retention priorities. 69

AWGN additive white Gaussian noise. 80

CA carrier aggregation. 4, 6, 20, 39–45, 51, 52, 55, 77, 78, 130

CAPEX capital expenditure. 47

CC component carrier. 21, 23, 40–46, 78, 79, 86, 128

CDF cumulative distribution function. 93, 94

CIF Carrier Indicator Field. 46

CoMP Coordinated MultiPoint. 3, 4, 6, 8, 19–22, 24, 26–32, 34–38, 61, 63, 67–69, 71–74, 76, 99, 102, 103, 115, 116, 124, 126–130

CQI channel quality indicator. 67–69, 127

CRS cell-specific reference signal. 57, 58

CS coordinated scheduling. 19, 22, 30, 32, 63, 67, 69, 74, 76, 127

CSG closed subscriber group. 49, 51, 53, 54

CSI Channel State Information. 45, 127

DC dual carrier. 78

DCI Downlink Control Information. 46

DL downlink. 59–61, 63, 66, 68, 71, 74

DRP downlink received power. 115–117, 120, 121, 123, 124

DSL digital subscriber line. 49

eICIC enhanced Inter-Cell Interference Coordination. 53, 54

eNB evolved Node B. 19, 21, 63, 66–70

FD Full-Duplex. 50

FDD Frequency Division Duplex. 40, 57

FDPS frequency domain packet scheduler. 69

FFR fractional frequency reuse. 55

FFT Fast Fourier Transform. 41

GBR Guaranteed-Bit-Rate. 42

HARQ Hybrid Automatic Repeat Request. 42

HD Half-Duplex. 50

HetNet Heterogeneous Network. 7, 9, 19–21, 47, 48, 50–53, 55–58, 129

HetNets Heterogeneous Networks. 4, 24, 130

HO Hand Over. 36, 43, 53, 57, 130

HSPA High Speed Packet Access. 78

ICI inter-cell interference. 3, 22, 23, 26, 53, 55, 59, 60, 63, 74, 126

ICIC Inter-Cell Interference Coordination. 49, 53–56

IoT Interference over Thermal. 54

JT joint transmission. 19, 22, 31, 32, 34, 61, 63, 67, 74, 102, 127, 130

LTE Long Term Evolution. 3, 7, 18, 19, 59–63, 66, 71, 74, 77, 78, 114

LTE-A LTE-Advanced. 4, 19, 20, 78, 126

MAC Medium Access Control. 22, 127

MCS modulation and coding scheme. 42, 64, 67, 68, 70, 76, 78, 127

MH Mobile Hashing. 23

MIMO Multi-Input-Multi-Output. 27, 130

MINLP Mixed Integer Non-Linear Problem. 23, 77, 97

MME Mobility Management Entity. 49, 50

MRC maximum ratio combining. 31, 70, 127

NLM network listening module. 54

OFDM orthogonal frequency division multiplexing. 61, 62

OFDMA orthogonal frequency division multiple access. 22, 60, 61

PA power allocation. 23, 72, 74

PDCCH Physical Downlink Control Channel. 46

PDCP Packet Data Convergence Protocol. 42

PF proportional fairness. 7, 22, 59, 61, 65–67, 71, 73, 74, 79, 127, 128

PRB physical resource block. 63, 64, 67–72, 76, 79

PS packet scheduler. 45, 46, 61, 67

PUSCH physical uplink shared channel. 31

QCI QoS Class Identifier. 42, 69

QoE Quality of Experience. 3, 18, 47, 50, 52, 59, 71, 78, 125

QoS Quality of Service. 19, 20, 22, 23, 42, 59, 60, 70, 77, 125, 128, 129

RA resource allocation. 78, 79, 128

RB Resource Block. 45, 86, 127, 128

RE range extension. 52–54, 56, 58

RF radio frequency. 41

RLC radio link control. 42

RR Round Robin. 23

RRC radio resource control. 42, 43

RRH remote radio head. 28, 29, 45

RRM Radio Resource Management. 20, 41, 77, 128, 130

RSRP Reference Signal Received Power. 56

S-GW Serving Gateway. 49

SINR signal-to-interference-and-noise ratio. 4, 19, 24, 27, 30, 32–35, 55, 58, 59, 63, 64, 67, 68, 70–72, 76, 80, 105, 125–127, 130

TDD Time Division Duplex. 40, 57, 58

TDM Time Division Multiplexing. 55

TDMA Time Division Multiple Access. 50

TDPS time domain packet scheduler. 68, 69

TP transmission point. 24, 103, 127, 130

TTI transmission time interval. 40, 46, 55, 62, 63, 65, 67, 69, 86

UE user equipment. 19, 63, 65–71

VoIP Voice over Internet Protocol. 42

Chapter 1

Introduction

1.1 Background and Motivation

The growth in the data traffic demands in the communication industry in the recent past has been explosive. The sheer number of subscribing users requiring full and reliable access to a range of highly data-hungry applications is what is driving such growth and overwhelm. High data rates are therefore needed to fulfil such mobile data-centric users' accessibility to these applications. With the current infrastructure of cellular networks, providing mobile and broadband connectivity to an increasing pool of users can compromise QoE. On the other hand, the existing trend spiral for consumption and reward is shifting towards higher provision of services at lower prices. Therefore, mobile operators are resorting to network-wide fundamental and infrastructural changes to boost revenue. To this end, new spectrally-efficient technologies that support the unprecedented outburst in the need for *true robustness* and data rates are needed for future deployment of mobile networks.

To meet the this objective, the International Telecommunications Union - Radio (ITU-R) issued performance specifications for 4G standards [1]. After LTE Release 8, as a 3GPP project failed to satisfy the requirement terms specified

by the ITU, the next-step evolved version of the LTE technology was released by 3GPP as the LTE-A. Major performance insufficiencies of the LTE [2] were in terms of required data rates, spectral efficiency and the support for variable bandwidth which LTE-A overcame.

Evolution-based innovations are adopted following the deployment of LTE-A further refining the spectrally-efficient techniques in LTE-A that were already existent in its predecessor LTE. Amongst such techniques that offer more enhanced utilisation of the scarce radio resources is the Coordinated MultiPoint (CoMP). Targeting interference-suffering cell-edge users with low average SINRs, CoMP techniques help increase user QoS significantly. Depending on the mode and level of coordination, exceptional spectral efficiencies are achieved by simultaneous transmission on the same resources. This is achieved by turning undesirable signals into ones contributing towards enhancing the SINR levels, in the case of joint transmission (JT). Same goal is obtained through the coordinated scheduling (CS) mode of CoMP by cooperatively scheduling transmissions amongst eNBs.

Moving towards 5G, *ultra dense* cell deployments have been shown to increase capacity in the entire cell beyond what is possible through conventional homogeneous networks. As a major capacity enhancing solution in LTE-A, Heterogeneous Networks (HetNets) have been proposed to favourably change cell dynamics. In contrast to single-tier homogeneous deployments, HetNets employ small cell eNBs to provide coverage for hotspot user equipments (UEs). Small cells also offer huge benefits in terms of spreading traffic loads as the macro base stations can offload a portion of the traffic onto the small eNBs. Additionally, in the cell-edge where robustness of transmissions from the high-power macro base stations cannot be guaranteed, existence of such cell can really pay off. Finally, given their low transmit power, small cells leverage a strong basis for capacity increase through spectrum reuse.

Projections on future volume for mobile data suggest that even operating at

theoretical bounds of information transfer, more bandwidth will still be needed to satisfy ubiquitous QoS demands. Acquisition of large bandwidth will not always be possible for operators. To this end, the flexibility of a network architecture to allow utilisation of smaller fragments of spectrum is very attractive. Introduced by the LTE-A, CA facilitates the simultaneous use of bandwidths of up to 100 MHz. With the evolution of many current link level solutions and technologies to the Shannon limit, importance of aggregating multiple carriers become ever more apparent. Although there are algorithmic and hardware challenges inherent to operating on multiple carriers, numerous flexible design options will arise as a result. Next section provides more details on the general objective of this thesis mentioned earlier.

1.2 Scope of the Dissertation

As briefly mentioned already, the main focus of this dissertation is the design of strategies and algorithms to enhance the capacity of cellular mobile communication systems. Both contributions made and research gaps identified have been derived on some of the most key pre-5G capacity-enabling technologies such as Coordinated MultiPoint, Carrier Aggregation and Heterogeneous Networks.

This Ph.D. thesis comprises of six chapters. The first main chapter is dedicated to providing a technical background into relevance and the motivation behind the research undertaken in this work. The key challenges for improving capacity with each of the technologies studied in this chapter are critically evaluated. Further, frameworks are established for potential candidate solutions in each case for obtaining the same macro objective of system capacity enhancement.

The main technical chapters 3, 4 and 5 are focused at providing strategies and algorithms for capacity improvement. This objective is pursued in chapters 3 and 4 from a radio resource management and packet scheduling perspective

whilst chapter 5 deals with the same goal from a cell association point of view.

Chapter 3 begins with addressing the detrimental effects of ICI on cell throughput in cellular networks reinforcing the significance of base station cooperation. Drawing upon the insufficiencies of some of the most relevant state-of-the-art technology available in the literature, it proposes a novel and comprehensive hybrid CoMP scheduling algorithm for an orthogonal frequency division multiple access (OFDMA) system. Its focus however, remains on reflecting on the joint benefits of the two modes of cooperation. In addition, full frequency reuse is identified as a tool to boost performance at the cost of elevated cell-edge interference. To this end, to facilitate the use of full reuse, intelligent base station cooperation is adopted on the Medium Access Control (MAC) layer followed by joint transmission in the physical layer to undermine the ICI downside in the cell-edge. The dual-eNB PF scheduling proposed for the JT and CS modes of CoMP provides a time and frequency priority basis for users. Moreover, the proposed scheme selects optimal transmission points (i.e. CoMP cluster) so to guarantee user minimum QoS. Through the assumed mobile network architecture based on C-RAN, implementation of coordinated scheduling is facilitated. Link adaptation is also a consideration in this work for improved spectral efficiency. A detailed algorithm structure is given following a mathematical representation of the objective it pursues. Key performance indicators for the algorithm are benchmarked against other cooperation modes of CoMP and results are analysed.

Resource allocation in a multi-carrier system is the research impetus in chapter 4. Non-contiguous carriers are jointly scheduled for transmission. The work addresses the joint problem of adaptive power and CC allocation. Having conducted a critical review of the related work, the direction for the research concentration is determined by the gap in the work existent in the literature. A paramount component to CC allocation is balancing the load on the carriers. Round Robin (RR) and Mobile Hashing (MH) are identified as the two principle

strategies for load balancing. Drawbacks are discussed for each method and a common cause is derived for such disadvantage as the inability of the schemes to guarantee instantaneous load balance hence a consideration in this research process. Having defined a tolerance factor for the load imbalance amongst the CCs, the scheduler's freedom in accessibility level to the resources available on the carriers is determined. Additionally, channel adaptive power allocation (PA) has also been studied in this chapter. The problem is formulated as a Mixed Integer Non-Linear Problem (MINLP) with maximising user sum rate as its objective function. Users' minimum QoS provisioning is also facilitated. The cast problem is solved in the dual domain using Lagrangian dual decomposition followed by the subgradient method. Analytical expressions are derived for the optimal power policy through which, the scheduler will be able to instantaneously guarantee CC load balance and user QoS. Finally, algorithm's performance is evaluated and analysed for various performance metrics under different imbalance tolerance factors and compared to baseline counterparts. The algorithm's capability in satisfying stringent QoS constraints is verified.

To fully reach the potential capacity limits in cellular networks, the importance of transitioning to denser network deployments are being increasingly recognised. Cell association dynamics in such heterogeneous environments are different to those in conventional homogeneous networks. Since cell-edge interference for small cells poses serious performance compromises, base station cooperation has been shown to interpolate several gains. Firstly, the ICI impact will not be as severe since interfering signals can be turned into desired ones. Secondly, cooperation facilitates higher SINR levels and this places more relaxed transmission power requirements, to achieve the same levels of SINR, on individual power-limited transmission points (TPs). Both of these factors are integral in achieving enhanced levels of capacity. This motivates the use of TP cooperation in conjunction with HetNets.

Chapter 5 shifts the focus of the previous chapters towards cell association with CoMP as the underlying technology for both UL and DL streams. To develop a cell association strategy for CoMP, a number of assumptions have been the put in place. Firstly, as with previous chapters, C-RAN is identified as a suitable architecture for the network deployment explained. In accordance to the literature review in chapter 5, assuming infinite capacity limits for the fronthaul data pipes between the base stations and the C-RAN is not practical or realistic. This becomes even more insensible when base station cooperation exhausts such links' real capacity limits. This calls for an association strategy that is aware of such limitation. Decoupled cell association in HetNets is studied in this work by assuming fronthaul link capacity constraints for CoMP. The decoupled UL/DL association strategy proposed builds up on the existing method for single transmission. It also ensures such fronthaul constraint is not violated taking into account multiple tiers of base stations. For improved fairness and performance, rate imbalance between the UL and DL is also a control factor as per the proposed scheme. Extensive simulations are carried out to compare the performance to other methods and comprehensive analysis of the results is provided.

Finally, chapter 6 provides concluding remarks on the entire research conducted and highlights the future improvements identified for individual phases.

1.3 Research Contributions

The contributions and novelties of this dissertation have been drawn from and are disseminated through the following technical papers:

[C1] A. Hooshmand, A. H. Aghvami, and B.W.K Ling, "Problem formulation for joint cooperative downlink scheduling and power allocation for joint processing coordinated multipoint," *in International Conference on Consumer Electronics*, IEEE, April 2014, pp. 1-3,

[C2] A. Hooshmand, A. Nallanathan, and A. H. Aghvami, "Joint inter-cell interference coordination and forced cooperative scheduling for the downlink of LTE systems," *in Wireless Communications and Networking Conference, WCNC*, IEEE, November 2014, pp. 1880-1884,

[J1] A. Hooshmand, H. Saki, and A. Hamid Aghvami, "Joint Adaptive Power Allocation and Resource Management for the Downlink of a Load-Balanced Multi-Carrier LTE HetNet," submitted to IEEE Transactions on Wireless Communications, August, 2015.

[J2] A. Hooshmand, M. A. Lema, M. Dohler, and A. Hamid Aghvami, "Fronthaul-Aware UL-DL Decoupling (FA-DUDe) for CoMP-enabled Heterogeneous Systems," submitted to Elsevier, International Journal of Computer and Telecommunications Networking, September, 2015.

Chapter 2

Background and Preliminaries

2.1 Coordinated MultiPoint

2.1.1 Motivation

To meet the unprecedented growth in mobile traffic volume, more efficient use of the available costly and scarce spectrum needs to be made [3] [4] [5] . However, exploiting the spatial re-use, techniques that offer tighter utilisation of the time and frequency resources introduce more ICI specially at the cell-edge. Therefore, cell-edge users are inherently more prone to link failures and heavily degraded spectral efficiencies due to the effects of inter-cell interference. The effect of ICI manifests itself on both the transmission reliability and speed while enhanced service reliability along with higher data rates are key impact factors for a more ubiquitous user quality of experience (QoE). Furthermore, increased mobile data volume is also driving the need for more densified cell deployments witnessed in recent years with heterogeneous nature. The ICI management becomes even more paramount in such scenarios where heterogeneous cell biasing is applied for load balancing. To this end, for a truly "*connected*" network, the essentially pivotal role of managing interference for such users residing at the edge of the network with low link budgets is exhibited more prominently. CoMP or Cooperative

Multi-Input-Multi-Output (MIMO) techniques will categorically address the issue of interference by either A) undermining its limiting effects through coordinating the transmission points (TPs) or B) exploiting it for improved cell-edge and average user data rates [6].

Obtaining higher spectral efficiencies requires high degrees of spectrum utilisation. In a full frequency reuse cell, a cell-edge user receives a similar SINR from its serving base station and the nearby interferers [6]. However, the effect of interference will also increase for higher reuse factors which makes network planning a vital function of the forthcoming mobile radio networks.

Starting out with the emergence of 2G, employing network planning to achieve spatial reuse lead to simple frequency reuse schemes. However, spectral efficiency gains brought forward by such schemes were low due to the spectrum underutalisation. Evolving into 3G and the Long Term Evolution (LTE), the more sophisticated inter-cell interference coordination (ICIC) techniques were introduced with full frequency reuse but limited gains as a results of increased inter cell interference, making them only adequate for medium network load scenarios [7]. ICIC schemes mainly rely on the exchange of messages between base stations on the X2 interface on a semi-static manner [8]. Such messages are to inform neighbouring base stations of the interference. Further down the evolutionary time-line of LTE beyond release 9 came the CoMP schemes with projected three-fold spectral efficiency enhancements for interference-limited networks. In practice of course, the gains brought by CoMP proved to be only in the 50% - 100% region in comparison to single point communication [7]. Later in this chapter, I will examine and identify factors affecting cooperative schemes in reaching their true potential.

2.1.2 CoMP Deployment Scenarios

Upon completion of a feasibility study undertaken by 3GPP in 2011, a work task commenced to introduce Release 11 specifications for CoMP in both uplink and downlink [9]. In general, coordination can be realised in two ways; 1) inter-site, in which multiple base stations can cooperate together in a geographical area to facilitate transmission or 2) intra-site, where multiple sectors of a single base station will cooperate. Remote radio heads (remote radio heads (RRHs)) or other self-sustained units can form such site sectors [6]. The X2 interface interconnects the base stations through fast fibre or a multi-hop link connection. There are four major possible deployment scenarios introduced by Release 11 in terms of transmission point cooperation with CoMP for both homogeneous and heterogeneous settings [8].

1. **Homogeneous intra-site macro connectivity:**

   Cooperation is facilitated amongst base stations at different sectors of a site controlled by a central base station. Such a cooperation requires no backhaul connection, making this mode of cooperation the least costly.

2. **Homogeneous inter-site macro cooperation:**

   Base stations involved in cooperation belong to different sites. The X2 interface will be utilised and therefore, there are high capacity, low latency and tight TP synchronisation requirements for the realisation of this base station cooperation deployment on the backhaul, specially for large CoMP cooperation sets.

3. **Heterogeneous cooperation with different cell ID:**

   In this case, low power transmission points within the coverage area of a macro base station will be cooperating and are each identified with their unique cell ID. In such a setting, orthogonal resource allocation ensures interference elimination between control channels and reference signals allocated. In general,

heterogeneous deployments of CoMP are more useful for metropolitan areas with high cell densities [9], enjoying a high cell-splitting gain that allows a large number of connections. However, the scheme's gain is somewhat compensated for by the limited choice of CoMP for resources on which the data channels of one cell collide with reference signals of another [8].

4. **Heterogeneous cooperation with same cell ID:**

   Cooperation in this case will be between a macro base station with other low power transmission points such as RRHs sharing the same macro cell ID with their base band units (BBUs) managed in a Cloud Radio Access Network (C-RAN). In this case, control channel and reference signal resources are shared between the transmission points. This will limit the control channel capacity making this mode of TP cooperation useful for lightly-loaded cell deployments [8]. Also, to have sufficient mutual information randomisation between either the demodulation reference signals (DMRS) and the CSI-RS that the user will need to receive in the downlink, it is important to ensure the allocation of different reference signal resources. Interference randomisation techniques such as the cell ID virtualisation introduced in Release 11, are employed to avoid interference on reference signal resources in the case of the use of the same resources [9].

2.1.3 User-Assisted Resource Management

In order to facilitate CoMP, user-measured link quality plays an integral role. A number of Channel State Information Reference Signal (CSI-RS) resources constitute a CoMP Measurement Set (CMS) for which users will measure the CSI related information such as the Rank Indicator (RI), Precoding Matrix Indicator (PMI) and the Channel Quality Indicator (CQI) [9]. Each of the resources on the CMS that are normally kept to a maximum of three to avoid overhead plateau,

will correspond to a single antenna port on a transmission point. Depending on the propagation characteristics of the terrain and the mobility scenario available, such measurements will need to be updated to ensure optimal performance of CoMP algorithms. The above-mentioned measurements are then transmitted by users on the PUCCH. The CoMP user procedure would therefore include user measurement of the CSI-RS, reporting the PUCCH containing such information to the network, reception of a Demodulation Reference Signal (DMRS) and then demodulation of the PUSCH to decode user data.

2.1.4 Uplink CoMP

Set to achieve both cell-edge and average user throughput gains of up to 300% and 80%, respectively, uplink CoMP exploits the previously detrimental interference signal effects and turns them into desired ones to boost SINR [6]. Uplink CoMP relies heavily on user measurements of the links both periodically or aperiodically. The Physical Uplink Control Channel (PUCCH) will contain control signalling necessary to implement uplink CoMP. It will carry scheduling request, (Negative)Acknowledgement (ACK/NACK) and CSI transmitted by the user to the network.

Although Release 11 leaves the uplink CoMP base station cooperation algorithms to implementation (i.e. vendors) [8], there are two major types of cooperation in the uplink, namely the joint reception and processing and the CS.

1. **Joint Reception (JR) and Processing:**

   The uplink JT scheme will enhance transmission throughput and reliability by accommodating for the same uplink signal to be received at multiple geographically separated reception points then combined at the final receiver. Forming a virtual antenna array, the Physical Uplink Shared Channel (PUSCH) transmitted by the user is received and processed by the receiving base stations.

The implementation complexity of this method is mostly due to the coordination of the reception points and the exchange of user data (i.e. quantised received signals) [9]. The pre-processing involved in joint reception can be flexible depending on operator needs and preferences.

On the reviver side, various pre-processed signal copies are applied maximum ratio combining (MRC) or Interference Rejection Combining (IRC) techniques such as Zero Forcing (ZF) and Minimum Mean Square Error (MMSE) to suppress interference and improve link quality. There are stringent backhaul capacity requirements for this technique as huge amounts of user and control data will need to traverse the X2 interface between the base stations. additionally, base stations will need to be tightly time and frequency synchronised to avoid inter-carrier interference (ICI) and inter-symbol interference (ISI). However, larger degrees of pre-processing will lower backhaul requirements but instead, decrease potential throughput gains and vice versa. Considering the low latency and high capacity backhaul requirements of the CoMP JT, centralising the baseband signalling units of the cooperating base stations in the C-RAN through fibre optic links is proven to be the only available mobile network architecture solution supporting such demands.

Another important feature of this scheme is the mismatch in the path loss compensation. This occurs when users receiving control signalling on PUCCH from a particular base station will be transmitting user data on physical uplink shared channel (PUSCH) to a different base station. Appropriate closed-loop power control commands are in place to correct such mismatch [9].

2. **Dynamic Point Selection (DPS):**

   For this less gain-achieving CoMP mode in comparison to the JT technique, multiple transmission points collaborate to optimally find a single receiver point with the most suitable interference profile with respect to the transmit-

ting user in the uplink. Scheduling information and decisions are exchanged between the base stations based on short term (i.e. every 1 millisecond) channel state information (CSI) reported by the users [9]. Less backhaul capacity is required to implement this technique although sensitivity is heightened to imperfect CSI due to becoming outdated.

2.1.5 Downlink CoMP

Base station coordination can significantly increase system capacity by eliminating or avoiding interference and hence achieving high SINRs [10] [9] [8]. In general, downlink cooperation scheme are far more standard-oriented than those designed for uplink. Up to the 3GPP Release 10 on the LTE roadmap, CoMP techniques had not yet matured sufficiently to be deployed. Simple coordination schemes appear on Release 11, however, further refining and enhancements are required for full realisation of CoMP for future releases. There are two major classes of coordination in the downlink distinguished by 3GPP, however, hybrid combinations of CoMP techniques can also be implemented for different interference scenarios. The classifications are based on the level of coordination between the transmission points and the backhaul requirements placed on the network as a result of the employment of such techniques. The two major categories of CoMP are namely CS or beamforming and JT [6] [9]:

1. **Coordinated Scheduling and Coordinated Beamforming:**

   User data is only available at a single TP with CS and CB. Users are scheduled for data transmission by different TPs in order for the received signal to have the highest possible SINR. Coordination amongst a CoMP set will determine a single TP to transmit user data. Such a transmission is configured by higher-layer radio resource control (RRC) signalling [9].

   (i) *Downlink Coordinated Beamforming (DL CB)*

Both transmitter-side and receiver-side beamforming algorithms function based on the superposition of waves principle. Transmitters employing beamforming will collaboratively control phase and amplitude of signals arriving at a receiver from multiple sites so constructive interference is achieved. Beamforming supported at the receiver-side however, operates on adjusting beam patterns so that a fixed gain is achieved towards the directions of the transmitters corresponding to each received signal. The output aggregate power of interference will be minimised [11]. In CB, in order to achieve this maximum SINR at the receiver, transmission power levels of the transmission points (TPs) involved as well as the beamforming coefficients and weights are calculated. This precoding achieves interference cancellation by creating interference nulls towards users scheduled by interfering TPs [9].

(ii) *Downlink Coordinated Scheduling (DL CS)*

In CS however, a central base station in a network cluster, with full knowledge of the interference situation of the neighbouring ones, will determine which radio links between users and other base stations will need to be active [12]. The feasibility and potential gains of this technique will highly depend on the X2 transmission technology employed, such as fibre optics or microwave connections and also the size of the cluster, as excessive overhead volume will render this scheme counter-productive. Dynamic Point Blanking (DPB) is an example CS-based algorithm in which interference arising from certain base stations are muted dynamically [9].

Due to the coordinated nature of CB and CS, base station cooperation needs to be facilitated through the exchange of CSI. Depending on given cell loads and available backhaul capacity and latency, appropriate feedback sharing mechanisms can be adopted. The CB and CS can be jointly implemented to achieve minimum multi-user and multi-cell interference [8].

2. **Joint Transmission (JT):**

With no standardised support in Release 11 for JT, simultaneous transmission of the same signal is arranged from multiple transmission points (TPs). In Long Term Evolution (LTE), the same resource blocks (RBs) are used for transmission to the same user in the downlink and therefore the suitability of JT is only limited to lightly-loaded networks [9]. This technique enhances the received SINR at the receiver as the interference signals transmitted from neighbouring base stations now become part of the final desired signal. This is particularly useful at the cell-edge of dense heterogeneous deployments in which, due to the vicinity of mobile users to other base stations, inter-cell interference effects place heavier performance burden on the network. Joint transmission (JT) requires copies of the same user data signals to be available at different sites. The scheme can be implemented coherently or non-coherently based on backhaul capabilities and operator requirements.

Backhaul requirements of this scheme is significantly higher than other CoMP techniques. Since the CSI are quantised and can be outdated very rapidly, JT remains sensitive to the errors in the reported CSI [7]. Fast changing propagation channel information between users and base stations need to be reported to base stations taking part in JT in order to dynamically perform link adaptation. Such frequent CSI reporting also avoids CSI out-dating which is best achieved through the joint base station scheduling. A joint scheduler's performance will be enhanced if signalling delay between base station is reduced. Baseband processing units (BBU) of multiple cooperating TPs can be grouped together in the C-RAN allowing for resource pooling and this way, low latency coordination can be made possible amongst base stations [8]. A fast backhaul solution such as a centralised BBU is particularly important to realise coherent JP in which precoding of transmission signals are carried out with consideration to those of other TPs' (i.e. spatial CSI feedback), requir-

ing better synchronisation and lower backhaul latency. In non-coherent JT, transmission signals are precoded individually with no knowledge about the neighbouring base stations' radio link quality with the users being served. Diversity gain is exploited in this mode of JT and strongest copy of the same signal arriving from different TPs is chosen and the rest discarded. Such a scheme will have the disadvantage of generating interference to other base stations as transmission power on transmitted downlink signals is higher to achieve better link quality [8].

LTE-advanced introduces Transmission Point Selection (TPS) as a special case of the JT mode in which users measure SINR received from base stations every transmission time interval (TTI) in a CoMP cooperation set [8]. The CSI corresponding to the best link along with the index of the base station achieving this maximum is reported. The signal is then transmitted in the downlink from this single base station. Users are therefore scheduled for transmission dynamically by the best base station. This mode of JT requires more frequent CSI reporting in comparison to the semi-static implementation of CS/CB and user data will only need to be available at a single site. Exchange of spatial characteristics and the realistic channel fluctuations facilitated by a fast joint scheduler is the major factor that help harness potential gains of TPS [9]. The above-explained modes of DL CoMP are illustrated in Figure (2.1).

As a newly 3GPP-presented solution for the next step of the current 3.5G mobile networks [56], LTE is perceived as a competent answer for recent boosts and urgencies in network services requirements [57]. Satisfying the growing needs for high data rates and QoS demands such as low latency and packet loss, all call for an optimised and totally revolutionised network with the potential to meet the above needs. The necessity of such high scale performance improvement becomes even more apparent by witnessing the 32-million LTE subscriber forecast by 2013 [58]. The above services are intended to be provisioned by an OFDMA

network in LTE [59]. Furthermore, amongst the functionalities that require algorithmic upgrade is the scheduling and resource allocation whose improvements will directly translate into enhanced throughput and capacity at the cell-edge. In this context, recent research is pursuing its direction towards the maximisation of network throughput, capacity and reduced delay in interference-bound cell scenarios in academia and industry.

To address the above challenge, identifying interference as one of the most preventative factors in throughput enhancement; ICI mitigating techniques such as frequency reuse have been studied in depth in the literature. In [60], given physical channel conditions and ICI, the performance of several frequency coordination schemes are evaluated for maximum cell throughput. It is shown that for different traffic types such as voice and data, no one particular frequency coordination scheme can have significant gains and that a trade-off needs to be found. It also introduces full frequency reuse as a performance-boosting scheme for data traffic. In the downlink (DL) optimal resource allocation decisions are derived mathematically in [61] in an effort to improve effective capacity gain whilst satisfying user QoS requirements. Also, such gains in the cell-centre are compared to those of the cell-edge. However, multipath fading as one of the main realistic and performance-degrading factors in resource allocation and scheduling seems to have been omitted in their considerations. The work in [62] focuses on analysing the performance of PF scheduling in presence of different network load. Closed-form expressions were also obtained which were supported by simulation results for throughput and fairness in OFDMA systems. Authors in [63] present a selective cooperation scheme in a two-cell cellular network in which system throughput is evaluated. Additionally a hybrid extension of the algorithm presented proves a superior performance in comparison to a full-time cooperation. However, the context in which the term "*cooperation*" is used in [63] only refers to JT and does not capture scheduling decision making between base stations. Ad-

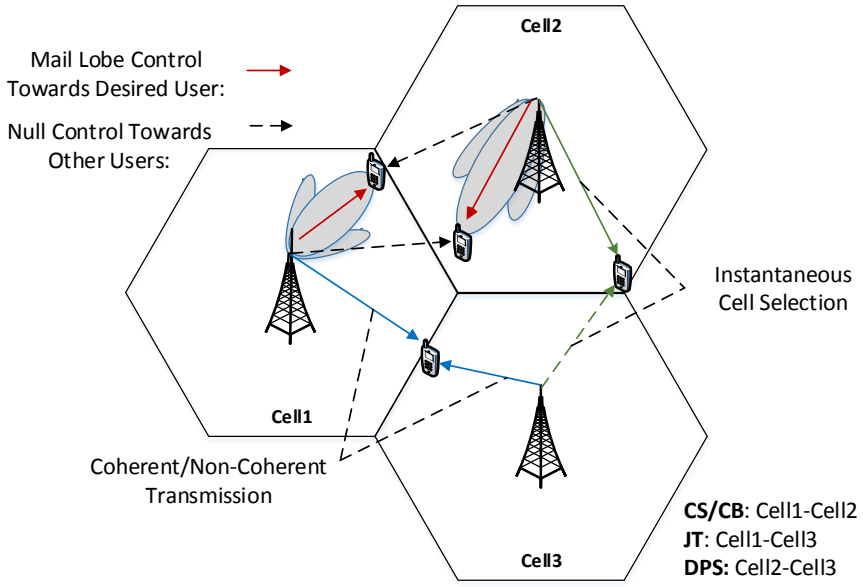

Figure 2.1: Base station cooperation modes in the DL of LTE.

ditionally, because cell-edge users are expected to have lower average throughput as a result of experiencing the worst average channel conditions, many research themes have been initiated to address the throughput maximisation of such users. Additionally, the underlying backhaul technology directly influences the choice of cooperation

2.1.6 CoMP Implementation Challenges

CoMP as a major enabler is envisioned for future mobile network generations. However, to reap the true performance gains, several limiting factors have been identified that require rectification. The extent to which performance measures of CoMP are compromised by any the following elements depend on the mode and deployment scenario at hand. Cooperation between the TPs in different sectors of the same base station does not require additional backhaul. Also the implementation of this mode of cooperation is relatively simple since the scheduling and processing functions required for CoMP can easily be upgraded and are currently co-located [8].

In the case of inter-site CoMP however, base station cooperation requires constant exchange of control signalling over the air interface and the backhaul. User measurements of the channel state information (CSI), preprocessed or quantised user data and user control data need to be shared amongst TPs synchronously over the X2 [6]. The long-term and short-term reporting of CSI are used for the purpose of Hand Over (HO) preparation and link adaptation, respectively. This will consequently require a backhaul capacity of around a few Mbps to 4 Gbps and approximate backhaul latencies of no more than around a few milliseconds. Large backhaul latencies will result in channel ageing [7] in which case, the outdating of the reported CSI will result in performance degradation of CoMP. To cater for this stringent backhaul requirement incurred by CoMP, base band processing unit (BBU) pooling offered by Cloud Radio Access Network (C-RAN) is identified as the most suitable candidate [13] [14]. Current backhaul latencies range from 100 Microseconds to about 20 milliseconds. Various factors that determine such delays can be either network technology related or stem from the choice of backhaul technology used. Processing delays introduced by network nodes and switching equipment are amongst examples of the former. The X2 backhaul connections such as conventional or millimetre microwave, multi-hop and fibre optic connections each contribute to latencies experienced in inter-site base station cooperation [6].

Another necessity for low backhaul latencies for CoMP to be feasibly realised is to avoid the limiting effect of such delays on the Hybrid Automatic Repeat Request (HARQ) that CoMP will need to be implemented with. To control the volume of feedback and control signalling traversing the X2, efficient feedback compression schemes need to be adopted specially in the case of JT CoMP with significant overhead and feedback associated. The CSI error is also introduced as another factor limiting performance gains of cooperation schemes that highly employ and rely on channel feedback. Channel prediction techniques at the pre-

coder side can reduce this dependency and sensitivity of such schemes (i.e. JT CoMP) to these errors [7].

Flexible clustering and user selection is identified as another solution for increased feedback [7] [8]. Although investigated to be NP-hard, the choice of base stations and users collaborating in CoMP will greatly impact the feedback required from users. Higher gains can be drawn from smaller clusters of cooperating base stations introducing less scheduling and processing complexities as the capacity loss due to a larger number of pilot sequences and increased feedback overhead can sometimes outweigh the capacity gains brought by CoMP through larger cluster sizes. Dividing the entire network area into small cells is shown to eliminate interference but raise the inter-cluster interference [7]. Dynamic clustering of TPs depend on the existing propagation environment and the user measurements of the radio frequency (RF) channels [6].

Moreover, the TPs taking part in CoMP are required to be time and frequency synchronised. Synchronising base stations on the frequency domain will eliminate inter-carrier interference (ICI). The inter-symbol (ISI) interference and the ICI effects are both mitigated through time-synchronous transmissions amongst TPs [6]. Further coherent reception of transmitted signals can be ensured through the recruitment of an effective TP phase-synchronising technique such as ones that function based on the Global Navigation Satellite System (GNSS) [8].

2.2 Carrier Aggregation

2.2.1 Background and Motivation

Users' expectations to receive high volume and reliable traffic data have shown an unprecedented growth in recent years and is projected to double every year in the current decade [15]. This is predominantly due to new and emerging data-hungry and personal hand-held devices such as tablets and smart phones [13]

narrowing user demands between mobile and fixed networks. To satisfy such user demands, International Mobile Telecommunication Advanced (IMT-A), a global standard initiative was introduced by the international telecommunications union (ITU) in 2007 [16]. The IMT-A requires peak DL and UL data rates of 1 Gbps and 500 Mbps for low mobility scenarios, respectively. This was when peak data rates of 300 Mbps and 75 Mbps at a maximum available bandwidth of 20 MHz were supported by the Long Term Evolution (LTE), corresponding to Releases 8 and 9 of the 3rd Generation Partnership Project (3GPP) [17]. Promising to enhance LTE's performance, LTE advanced (LTE-A) soon was issued as an IMT-A technology two years after its introduction by 3GPP in 2010.

Since reachable data rates increase linearly with bandwidth, acquiring more spectrum is a necessity for meeting the ever-growing traffic requirements. LTE-A allows the utilisation of a maximum of 100 MHz system bandwidth. However, due to unavailability of large fragments of contiguous bandwidth, operators seek alternatives to use spectrum chunks at different carrier frequencies and aggregate them for data transmission. First standardised in Release 10 as one of the key features of 3GPP, CA facilitates the aggregation of fragmented and non-contiguous bandwidth as an expensive and scarce commodity [18] [19] [20]. This thesis only considers CA in the DL.

2.2.2 Aggregation Modes and Cases

Utilising the bandwidth fragments or otherwise known as CCs can be done in three ways depending where they lie on the spectrum. Available in both Frequency and Time Division Duplexing (FDD and TDD) modes, in all of the LTE-A CA modes, CC allocations follow that of the LTE with possible CC bandwidths of 1.4, 3, 5, 10, 15 or 20 MHz. The Frequency Division Duplex (FDD) CA, as the only CA mode considered in this thesis for the DL, may allocate different CCs for uplink and downlink transmissions [20] [19] whereas in Time Division Du-

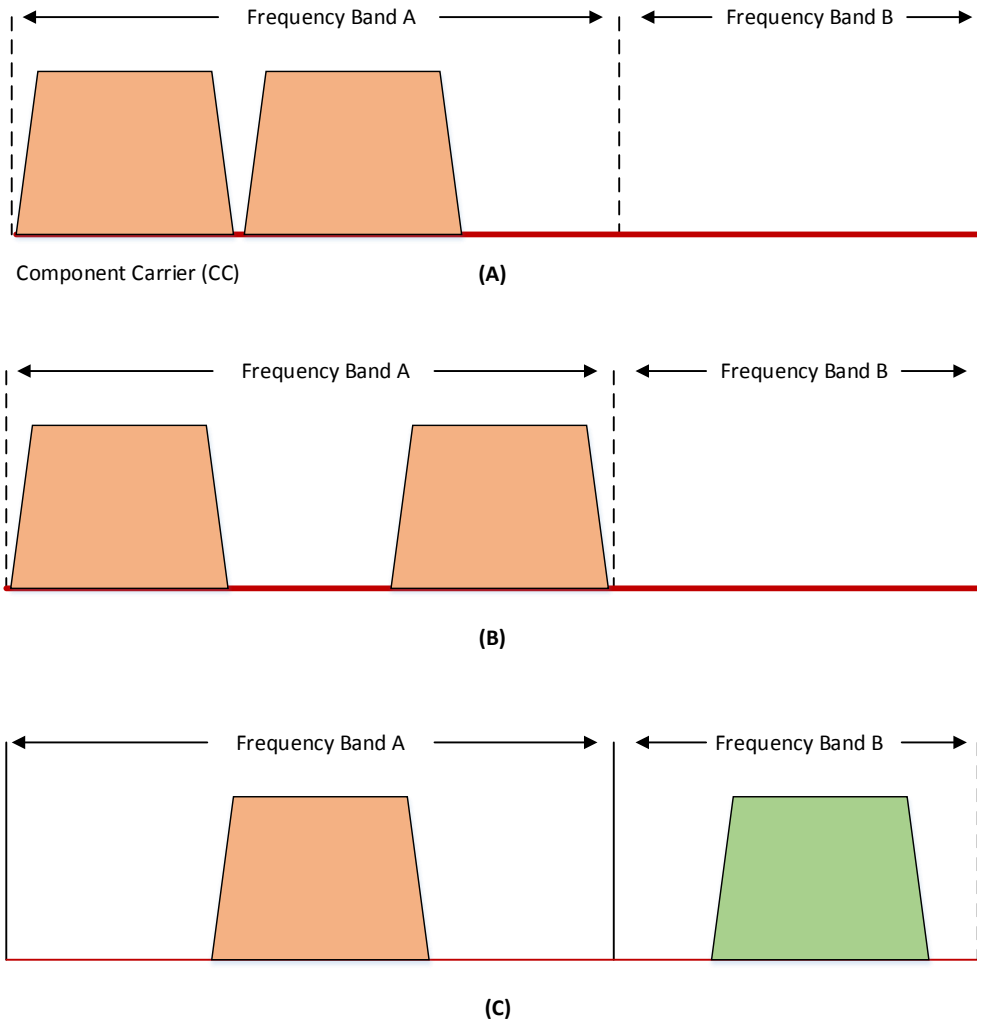

Figure 2.2: CA in LTE-A. (A): Intra-band contiguous aggregation, (B): Intra-band non-contiguous aggregation and (C): Inter-band non-contiguous aggregation

plex (TDD), uplink and downlink transmissions are scheduled at different LTE subframes or transmission time intervals (TTIs) of length 1 millisecond over a single CC [17]. An aggregation of up to 5 CCs is facilitated in LTE-A, all of which can be managed by a single Mac layer entity [17]. As illustrated by Figure (2.2), there are three modes of CA through which operators can create data transmission pipes out of the available CCs:

(i) *Intra-Band Contiguous*

In this mode of CA in LTE-A, adjacent CCs are used for data transmission [18], Centre frequencies of contiguous CCs will be allocated in accordance to a spacing of 300 KHz to maintain compatibility with 100 KHz spacings of Release 8. Majority of current frequency allocations today do not allow for very wide contiguous allocations however, at relatively high carrier frequencies (i.e. 3.5 GHz), intra-band contiguous CA might be a feasible operator option [20].

(ii) *Intra-Band Non-Contiguous*

The allocation of frequency-separated CCs results in a non-contiguous allocation. However, the spectrum chunks will still be in the same band in this mode of CA [19]. The intra-band non-contiguous allocation serves as a suitable frequency allocation solution for operators that support network sharing and have the middle CCs occupied to other users [20].

(iii) *Inter-Band Non-Contiguous*

The scattered bandwidth segments acquired by operators at different bands are facilitated for transmission using inter-band non-contiguous CA. Introducing additional radio frequency complexities and some user capability requirements, this mode of CA exploits the radio propagation characteristics at different bands offering a new level of mobility robustness to users [17] [20].

The transition from the Release 10 of the LTE-A in which only a limited number of combination of bandwidth aggregation was possible to Release 11 with

full support for non-contiguous was rather easy as only the radio frequency (RF) requirements needed to be specified [18].

2.2.3 Implementation Considerations

The above CA modes are designed to incur minimal burden on the existing network infrastructure and elements. Here, I investigate some of the main considerations and design objectives for incorporating CA into the future releases.

The feasibility of CA in UL is not always as straightforward as the DL case. The overhead burden on power-limited terminals in the UL due to CA makes the gains achievable less promising. In addition to the elevated terminal power consumption and increased control signalling that arise from utilising multiple CCs, there are other factors that network operators will need to account for. Considering the different radio channel characteristics of CCs at different carrier frequencies and hence different path losses and Doppler shifts, non-contiguous CA in LTE-A requires RRM and radio network planning algorithms to account for such differences [21]. Non-contiguous allocation of CCs additionally requires multiple RF and Fast Fourier Transform (FFT) units [18].

Per CC Allocations:

RRM functionalities between the multi-carrier enabled LTE-A and the Releases 8 are mostly similar to ensure backward compatibility for the previous release legacy users' accessibility. For example, admission control in both releases is done at the base station prior to the establishment of a new CC. Furthermore, each CC (i.e. radio bearer) requires a unique Packet Data Convergence Protocol (PDCP), radio link control (RLC), transport channel, Hybrid Automatic Repeat Request (HARQ) entity and separate power levels to control CC coverage [22] [23]. Link adaptation is also carried out individually for each CC on the independent transport blocks with individual modulation and coding schemes (MCSs). However, CC Configuration is a new base station vendor-specific functionality that

only the LTE-A CA-enabled is equipped with [18]. CC Configuration configures CCs to each user through radio resource control (RRC) signalling. A number of factors determine the quantity of CCs allocated to each user. The QoS parameters of a user such as the QoS Class Identifier (QCI) for different traffic types along with terminal capability to support CA are amongst those considered [24]. Priority Guaranteed-Bit-Rate (GBR) users are often allocated multiple CCs, in contrast to the Voice over Internet Protocol (VoIP) ones, to ensure their QoS requirements are met.

Necessary Upgrades on the Control Side:

To serve as a backward compatible solution to the legacy users of early releases of the LTE, CA in LTE-A has been designed to require minimal upgrades on the procedures and functionalities already existing within the LTE's framework. The RRC for instance is identical in an LTE-A based CA system to that of Releases 8. This per user functionality facilitates users' reception of radio related information from the network. The RRC_IDLE_User procedure is not affected in LTE-A in which idle users that cannot transmit or recieve data to/from the network become connected users through the establishment of an RRC connection. On the other hand however, cell management and cell activation/de-activation capabilities become paramount to a CA-enabled network. Here, a CC, a radio bearer and a serving cell are used interchangeably, as they are treated the same by the higher layers [17].

Cell Management:

With CC selection being a user-specific precedure, each user can be configured to more than a single CC for data transmission in LTE-A with CA. Such CCs will therein be considered as the user's primary or secondary cells depending on whether or not they provide the user with control information and functions as well as data transmission [24]. Control signalling will only be sent to a user through a single primary cell (i.e. PCell). Data transmission however, can occur

through a user's PCell and secondary cells (i.e. SCells) [18] [20]. Cell management is an RRC-based control procedure to allow for the removal and change of such cells by the network, providing a useful tool for the network to optimise the users' power consumption. Not assuming HO, changing PCells is not possible for any user in th network as it is a permanent connection requirement for the network for control signalling [17].

Cell Activation and De-Activation

Another key feature that LTE-A CA supports is the ability of the network to activate or de-activate some of the cells a user is configured to. The main purpose of this task is to prolong terminals residual battery. Again, as explained above, PCells cannot be de-activated due to the provision of control signalling [20].

Necessary User Enhancements:

Through a high degree of similarity in the terminal capabilities of legacy users of the LTE and LTE-A with CA, backward compatibility of CA as an enabler is established making it an integral component of the future releases. On the other hand, Releases 10 and 11 of LTE-A have made bandwidth utilisation more spectrally efficient through emphasising on terminal upgrades. To this end, LTE-A users exhibit capabilities to allow for CA in addition to inheriting those of the predecessors'. Transmitting and receiving control channel signalling simultaneously in a single PCell is amongst these upgrades. Also, transmission of data from multiple antennas with spatial multiplexing is another key LTE-A user enhancement [17].

Performance Boost Through Overhead Reduction:

As discussed earlier, employing more number of CCs will result in higher overhead throughout the network. Current CA schemes utilise compatible bandwidth fragments over which the transmission of the broadcast, reference, synchronisation and downlink control channel signals are carried out. Such overhead is irreducible and inevitable regardless of the quantity of CCs used [17]. Practically,

the use of small bandwidth fragments in certain network scenarios makes CA not worth it. In order to widen the gap between the performance gains of CA and the loss due to heavy backhaul demand because of the overhead, introduction of non-compatible user-data-only CCs proves effective in future releases.

2.2.4 Deployment Scenarios

Exploiting different radio frequency (RF) propagation dynamics at different carrier frequencies, CA is designed to boost data rates in overlapping cell areas [17]. Generally, higher carrier frequencies used for data transmission will have less coverage as they attenuate more. The following are the five identified cases under which CA can be implemented with the assumption of having two bandwidth segments with carrier frequencies at F1 and F2 where F2 >F1 [25].

1. **Intra-Band Co-located and Overlaid Cells**

   The cells in this case will reside on the same band with almost the same coverage. Data rates are enhanced through the aggregation of multiple bands. Users in the entire cell area can benefit from higher achievable throughputs.

2. **Inter-Band Co-located and Overlaid Cells**

   Since F2 >F1, therefore, the path loss associated with transmissions on F2 will be higher, yielding less coverage. F1 will be used to support mobility and coverage and F2 employed to further increase throughput.

3. **Inter-Band Co-located Cells with Directed Antennas**

   Existing coverage holes due to the larger path loss are covered by directing the antennas of the higher frequency carrier (i.e. F2) towards the cell boundaries of the component carrier with less attenuation (i.e. F1).

4. **Inter-Band Cells with HetNet-Type Coverage**

All users within the cell will be within the macro coverage of the lower frequency band F1 and the short-radius coverage F2 bands are recruited by RRHs to provide higher hotspot throughput.

5. **Inter-Band Co-located and Overlaid Cells with Repeaters**

   Similar to case 2, different bands are used but in conjunction with repeaters to extend the coverage of the available frequency bands

2.2.5 Packet Scheduler of LTE-A with CA

The layer 2 scheduler of the LTE-A is identical to that of Release 8 in that it uses the same frame structure and resource allocation granularity as the LTE counterpart with access to Resource Blocks (RBs) of 180 KHz bandwidth consisting of 12 sub-carriers. Due to multi-carrier capabilities needed in Release 11 however, extra design enhancements have been incorporated to facilitate CA. Cell activation and de-activation is one of the add-on procedures to the LTE-A packet scheduler (PS). With the main objective of controlling terminal power consumption, it de-activates users' CCs if no data transmission is scheduler for a user on a particular CC within a pre-specified time period. This timer-based behaviour ensures no Channel State Information (CSI) is reported by the users on de-activated CCs and hence prolong batter lifetime [18].

LTE vs LTE-A Scheduling Grant:

Another element of the LTE-A PS is the cross-CC scheduling feature in which the scheduling grant by the base station may be transmitted on a particular CC for transmission on another. Users' CC allocation information, otherwise known as Carrier Indicator Field (CIF) is appended in the Downlink Control Information (DCI) section of the Physical Downlink Control Channel (PDCCH) and is signalled to users before data transmission. Release 8 scheduler on the other hand, transmitted the scheduling grants on the CC-specific and time-multiplexed

PDCCH each TTI with the same addressing for each user no matter what CC allocation had been configured to that user [18]. Finally, an optimal PS considers CC load scenarios (i.e. own cell loads) for optimal system performance [26].

Carrier Load Balancing:

Irrespective of whether or not a single PS manages all available CCs, the CC selection will need to factor in the potential impact of the carrier load balancing on system performance. Joint scheduler schemes assume a single entity to carry out the scheduling tasks of CCs whilst the independent scheduling generally allocates separate schedulers to each CC [27]. There are two classes of CC load balancing algorithms in the literature: The Least Load (LL) Balancing or also know as the Round Robin (RR) and the Mobile Hashing (MH) or the random carrier balancing [28] [29]. The former ensures load balancing through scheduling of new user onto CCs with least amount of load whereas the latter maps the hash output values of user terminals hashing algorithms to the indices of CCs. Moreover, terminal capabilities also significantly influence scheduler flexibility in load balancing.

Driven by the need for increased user data, QoS-driven resource allocation (RA) is seen as an integral design feature of the forthcoming 5G infrastructure in order to guarantee a minimum QoE for users. On the other hand, with current wireless links approaching their theoretical capacity limits, the urge for acquiring more bandwidth becomes more pressing. However, often times, wide bandwidth fragments are not available for operators, therefore, aggregating smaller chunks will help enhance data transmission for the data-hungry applications and users. carrier aggregation as a technology that makes this task possible was first introduced in High Speed Packet Access (HSPA) cellular systems as a dual carrier (DC) access scheme. However, only aggregation of adjacent and contiguous carriers in the DL and UL was possible [76]. LTE-A however, facilitates the aggregation of non-contiguous spectrum belonging to different frequency bands [77].

Round Robin (RR) and Mobile Hashing (MH) as the two well-known carrier

load balancing schemes are thoroughly investigated in [78] and compared. Cell throughput of users and coverage are presented under both algorithms for different queue types and traffic models in this work. Authors investigate different cases of packet scheduler, namely *per-CC* PS or a central PS overseeing the task of resource allocation for all available CCs. However, as addressed in the work, both algorithms can only offer long term fairness amongst the CCs employed and hence a consequent reduction in trunking efficiency. Study undertaken in [79] proposes dynamic CC allocation. Authors present a sub-optimal greedy algorithm for a network with CA considering the MCS for the RA problem. Since power allocation is not dynamic, the algorithm's efficiency is limited and only comparable to those with equal power allocation. Carrier load balancing is also not a consideration here. A similar equal-power physical resource block (PRB) and CC allocation algorithm is also presented in [80] achieving a long-term PF in RA amongst the users. Authors in [81] propose a sub-optimal RA algorithm with adaptive power allocation and solve the problem using a standard decomposition technique. However, minimum QoS considerations are omitted in their analysis. Carrier load balancing also has not been investigated in this work. In [82], a fixed power-level iterative and decentralized RA algorithm is presented. The algorithm maximises the logarithmic and sigmoidal-like utility functions guaranteeing a minimum data rate for the multi-carrier system. Proof showing the single-multi RA equivalence under certain scenario is also presented. Further, a novel price-based resource allocation problem is cast in [83] under elastic and inelastic user traffic. Authors show that an optimal solution to the equal-power algorithm presented is tractable by proving convexity. However, with dynamic allocation of power and carrier load balancing assumptions omitted, the results obtained in this work remain limited. In addition to the above, power allocation resource management schemes with a finite power levels are also studied in the literature.

2.3 Heterogeneous Networks

2.3.1 Background

The continuous explosion in mobile wireless data requirements in the recent years call for feasible and adequate treatments. Enhancing the efficiency of the current technologies no longer seem sufficient in maintaining the supply to demand ratio as the already efficient wireless links are reaching their theoretical bounds. Also, with the increase in the number of subscribers and smart devices specially in the dense urban population areas, the importance of cell splitting and employing more bandwidth becomes more visible. To meet such expectations for a more uniform user QoE, higher spectral efficiency gains are required per unit area from the future networks [30]. High density base station deployments, as the new network topology have been proposed to offer higher quality links in addition to orders of magnitude higher spectral efficiency gains of up to 1600 fold [31]. Reduced transmitter-receiver distance will have dramatic affect on link robustness and an increased level of spatial reuse gain. This is achieved by reducing cell sizes through a heterogeneous deployment of base stations [32] [33]. The HetNet architectures as an alternative to the current homogeneous deployments, present less capital expenditure (CAPEX) to the operators coupled with less cost per bit transferred through the wireless network [34]. Moreover, the deployment of HetNets is reported [35] to have reduced the network costs by 70% in comparison to a homogeneous strategy. The low power nodes of such networks also alleviate macro coverage holes and increase capacity in hotspots nearby where more resources are required with poor quality macro wireless links. Further, the requirement of acquiring additional expensive urban-area macro sites are also bypassed through the deployment of HetNet strategy [30].

2.3.2 HetNet Nodes

Distinguished broadly by their transmit power level and backhaul are the small cell nodes in HetNets [36]. The underlying macro base station providing full coverage for such small cells will operate on approximately 5 - 40 W transmit power. Local to any hotspots, the small cells provide a reliable service in addition to alleviating macro coverage holes. High degree of flexibility is offered by the low power nodes as they can have access technologies different to that of the macro base station (Cellular, WLAN, etc) [37] [38]. Costs associated with small cells are also significantly lower that macro base stations. Outdoor low power nodes operating between 250 mW - 2 W of transmit powers do not need air conditioning units for their power amplifiers, proving to be even more cost effective.

Next I present the low power base stations considered as transmission nodes in HetNets [34] as depicted in Figure (2.3) and discuss some of their key features:

1. *Pico Cells:* Similar to regular macro base stations from an architectural perspective, Pico cells have lower transmission power ranges. They are limited to 100 mW and between 250 mW - 2 W for indoor and outdoor deployments, respectively. Pico base stations require network planning and have X2 backhaul connections and hence facilitate the implementation of Inter-Cell Interference Coordination (ICIC) techniques. There are both fibre optic wired and wireless backhaul connections defined for Pico cells. The former is costly and offer rates of up to 50 Mbps. The wireless backhaul connectivity can be implemented on both licensed 6-42, 70-90 GHz or unlicensed 2.4, 5 and 60 GHz carrier frequencies [32]. Similar to macro base stations, Pico cells connect to the Mobility Management Entity (MME) and the Serving Gateway (S-GW) through the S1 interface.

2. *Femto Deployments:* Lowest class of transmit power level of a maximum of 100 mW belongs to Femto or otherwise called Home Base Stations (or HeNBs)

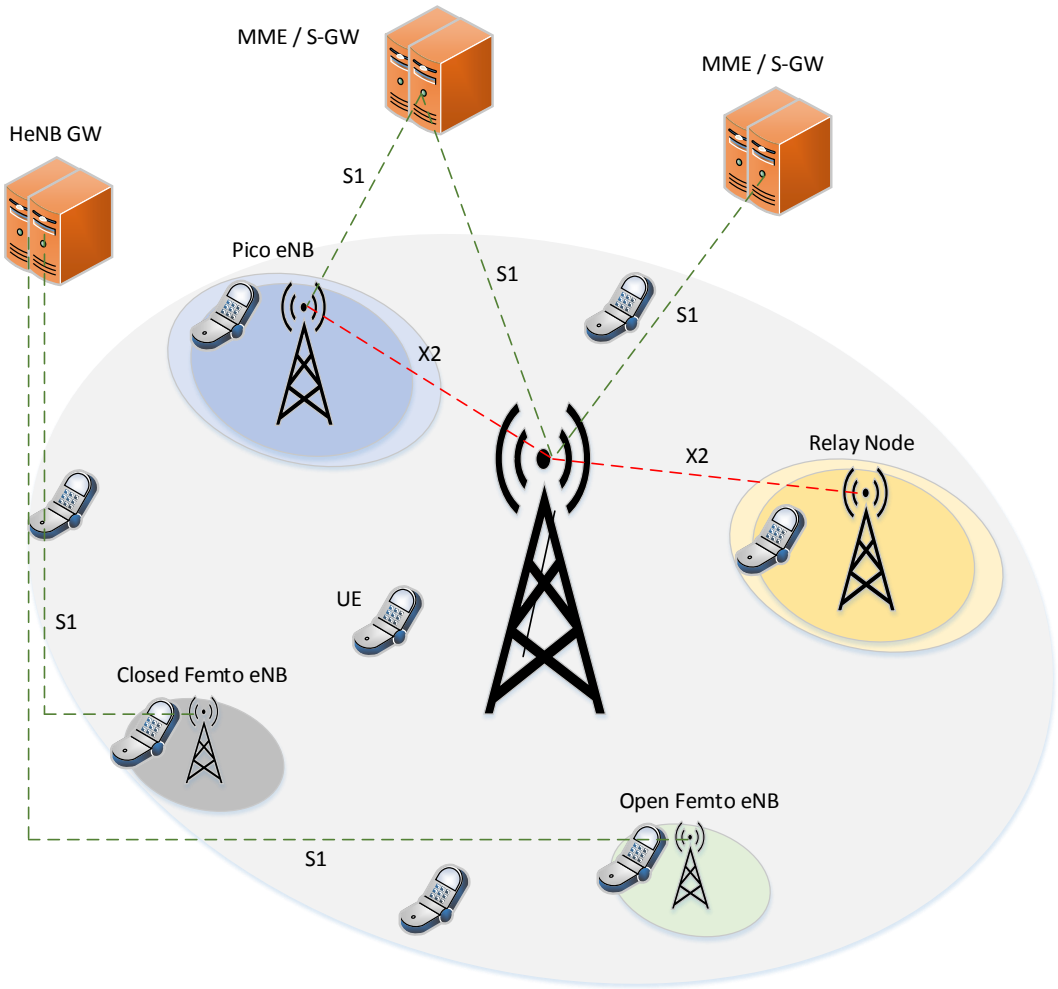

Figure 2.3: A HetNet consisting of different transmission points with various transmit power levels.

equipped with omni-directional antennas that only support unplanned indoor access. Femto owners can configure open or closed access for such nodes. They can use consumer home digital subscriber line (DSL) or cable modem for network backhaul connection. While open-Femto provides accessibility to all local users, a closed-Femto only grants access to users belonging to its closed subscriber group (CSG). This type of Femto cell creates macro coverage holes and outage regions proportional to its transmission power levels. With no X2 interface, ICIC schemes in LTE Releases 8 and 9 are not supported for closed-Femto deployments, in case of co-channel deployment with the underlying macro base station. Additionally, the static OAM-based techniques available in LTE-A Release 10 only rely on transmit power control.

3. *Relay Nodes:* Wireless backhaul connects relay nodes to the rest of the network when there are air interface resources available [39]. Designated for both indoor and outdoor scenarios, relay nodes transmit power levels are limited accordingly. Transmission power regulations limit relay nodes to a maximum of 100 mW for indoor and between 250 mW - 2 W for outdoor communications. Directional and omni-directional antennas are employed for backhaul and access link connectivity, respectively. There are two modes of relay node deployment strategies based on their resource utilisation pattern. As the main 3GPP focus, In-Band relays use the same radio frequency resources in both uplink and downlink for the wireless backhaul communication. Out-of-Band relays however, use different resources requiring dedicated spectrum. Both Full-Duplex (FD) and Half-Duplex (HD) communications are further defined for In-Band relay nodes. Simultaneous transmission and reception is facilitated for FD in which the same frequency band is used on the relay to base station and user to relay radio links as opposed to a Time Division Multiple Access (TDMA) type communication in HD. Relay nodes have the same MME connection interface as pico and macro cells.

2.3.3 Deployment Strategies

HetNets offer substantial performance gains through reducing cell sizes. High traffic demands of hotspots and urban areas can specially benefit such dynamics in meeting a uniform user QoE. In this section, I present possible strategies under which potential benefits of HetNets can be exploited:

1. *Co-Channel Deployment:*

   Employing radio resources residing at different frequency bands will need user terminals to support aggregation of carriers. On the other hand, it will introduce further complexities in terms of cost of acquiring additional bandwidth and accounting for different propagation characteristics of different bands. Co-channel deployment of HetNets avoids segmentation of bandwidth layers and hence avoid many of the downsides specially for operators that do not have the infrastructural means to undergo such costs. HetNets provide a dynamic in which both the number of interferers and interference magnitude increase as a result of the increased network density. However, at the same time, useful signal strength from intended transmitters also rise as a result [40]. The coverage complications that occur with co-channel deployment in the context of closed-Femto cells call for intelligent management of interference caused by the low power nodes to users attached to the macro base station. Such users within the vicinity of the privately deployed Femto nodes are denied access leading to deep coverage holes. Likewise, users served by small cells (i.e. relay nodes, Pico or Femto) also experience heavy interference due to the strong macro coverage over their local connections [34].

2. *Multi-Channel HetNets:*

   A potential solution to rectify the above-mentioned problem of coverage holes caused by closed-Femto that are only accessible to their CSG users, multi-carrier access for HetNets proposes to allocate different frequency bands to

small cell and macro users. Although this scheme requires additional spectrum, in a special case of the multi-carrier deployment in which other small cells are used in conjunction with open-Femto cells, users with CA capabilities will specially reap the performance gains accrued by this technique [41] [34].

3. *CA-enabled HetNets:*

   With time, as mobile terminals evolve gradually to support aggregating multiple carrier, operators will have higher flexibility to utilise bandwidth fragments across the spectrum. HetNet deployment can particularly take advantage of CA to protect both macro and closed-Femto cell users. One of the most attractive deployment scenarios involving CA in HetNets is allocating a dedicated band to the users of the underlying macro base station at full transmit power for coverage to protect such users from the closed-Femto interference. A second band can then be considered for all users in the coverage area of both the macro and the closed-Femto cell, whilst the macro will be transmitting at a fraction of its maximum power on this band to reduce disparity of power with the small cell [34]. This scheme will help overcome the under-utilisation problem of small cells by driving more load towards them as users will not see much difference in link quality between the macro and closed-Femto [42].

   In the UL effective solutions such as the *Dual Connectivity* also known as the inter-site CA initially introduced by 3GPP in [43] have been investigated to increase the *per-user* throughput. Inter-site CA allows users to have access to the resources across the macro and small cells.

2.3.4 Motivation for Interference Management

Limited by their coverage area due to their low transmit power, small cells in HetNets are not exploited sufficiently for their potential to enhance network capacity [44] [45]. This is simply because of the high power disparity between these

nodes and the macro base stations. The adverse effect of small cells not drawing enough traffic is practically double-edged since macro users also suffer the unavailability of sufficient resources. To this end, one can easily conclude that a mechanism needs to be in place to balance the load amongst macro and small cell base stations in a HetNet for an optimal low power cell utilisation.

The concept of range extension (RE) first initiated for LTE, has been proposed to balance the load between TPs in a HetNet [32]. Both cell splitting gain and the user QoE due to a more balanced load, are significantly increased through RE. However, optimal cell association is victimised as a result, as not all users will be connected to the base station with the highest link quality [46]. To extend the coverage of small cells, a positive bias is considered. When the difference in signal strength measured by the user terminal from the macro and nearby small node falls below this bias value, an association to the small cell is triggered, adding higher radius coverage contours around the low power nodes. Therefore, the bias values considered dictate the HO boundaries. It is readily deduced that enhanced performance in HetNets owes itself to determining optimal HO boundaries to avoid small cell radio link failure, in co-channel deployments [34]. Consequently, users in the extended coverage areas suffer strong macro interference affecting the robustness of data and control signal transmission links to these users.

To allow for RE, macro interference can be handled either through terminal interference cancellation or base station coordination [30]. The main purpose of enhanced Inter-Cell Interference Coordination (eICIC) techniques is to leverage cell splitting gains through RE in HetNets weeding out the interference sources.

2.3.5 Interference Management

Co-channel implementation of HetNets create dynamics in which interference becomes the network performance paralysis reason. In this section I discuss some of the interference scenarios that arise in HetNets and then classify the existing ICIC

techniques in accordance to the needs stemming from the possible deployment strategies. As mentioned in the previous section, users as well as base stations can take part in diminishing the effects of interference. Resource partitioning has been identified as one of the most effective tools in base station resource coordination to eliminate ICI. This scheme allows non-CSG users to recieve data when near closed-Femto base stations. Scheduling some resources with high transmit power will increase coverage whereas resources with low transmission power increase capacity through high cell split gains [30]. Moreover, resource partitioning configures interference-protected resources and this mitigates UL interference as a result [34].

2.3.5.1 Interference Scenarios

(i) *Closed-Femto Deployment:* Coverage blind spots are created in the DL for macro users in HetNets with closed-Femto deployments because of the restricted association policy enforced [44]. To a greater extent, non-CSG users of the macro base station will undergo severe interference due to their proximity to the closed-Femto node.

To address this, the DL transmission power control techniques such as the *Dynamic Femto Cell Power Control* as introduced in [47] [48] offers remarkable reduction in interference. The objective of this policy is to prevent any interference leaks to outside of the Femtocell coverage area. A tolerance ratio for the Femto to macro power is defined for the non-CSG users. Femto cell transmit power is intelligently adapted so not to exceed this ratio by the means of monitoring macro signal power using a network listening module (NLM).

(ii) *Open-Acess Small Cell Deploymemt:* Interference can occur if open-access small cells such as Pico cells and relay nodes use the same time-frequency resources as the macro base station in the coverage-extended areas due to RE. To avoid this, Release 10 supports the X2-based ICIC technique of *Adaptive*

Resource Partitioning. Available only to cells with X2 connectivity (i.e. Pico cells), this policy uses a subframe as its resource allocation granularity to allow cell biasing. The Almost Blank Subframes (ABSs) not used by the macro base station, as the dominant interference source, will be declared to the Pico cell for potential use [34].

2.3.5.2 eICIC Classes

Interference management schemes in 4G allow for resource partitioning. Additionally, eICIC techniques are also required to allow flexible cell associations for reduced UL/DL interference [49] [30]. ICIC techniques in LTE Release 8 can be categorically grouped into two branches. Reactive methods respond to governing interference dynamics in the network whilst strict Interference over Thermal (IoT) values will be maintained in the UL [34].

Full frequency reuse has been investigated to cause unprecedented amounts of interference in HetNets. Therefore, fractional frequency reuse (FFR) techniques were proposed as part of the Proactive ICIC for LTE [32]. There are two types of FFR ICIC. The simple-to-implement strict FFR eliminates inter-cell interference by reusing frequency resources in inner and outer regions in a partitioned manner based on SINR. Same frequency bands are used for cell-centre users in every cell. Soft FFR however, reuses the entire available band in all cells with different power variations. Cell-edge users enjoy higher transmit powers to ensure acceptable link qualities.

The above ICIC techniques considered prior to the development of LTE-A also used Time Division Multiplexing (TDM) are refined in Release 10 onwards to facilitate CA in managing ICI. Advanced interference management methods fall in two major categories [30]:

1. *ICIC:* As above, ICIC techniques are required to mitigate interference on both data and control channels. Aware of cell load and the quantity of the deployed

small cells, such techniques offer resource partitioning to increase cell splitting gain. There are three tools that ICIC techniques can benefit from: *Resource coordination* as used in resource partitioning in time, frequency and space, *power control* as in closed-Femto transmit power control and *spatial beams adjustment*. This type of interference management is carried out on a TTI basis. Very tight base station synchronisation over the X2 is a strict requirement for these techniques.

2. *Slowly Adaptive Interference Management:* Over much larger intervals, the central scheduling entity of this class of interference management technique allocates resources in an attempt to determine optimal user and base station transmission power levels. Optimal calculation of such parameters over all frequency resources will generally take longer, therefore similar schemes might not be practical in many network scenarios. With high bandwidth requirements and computational complexity, slowly adaptive interference coordination techniques are more desirably implemented in a distributed way.

2.3.6 DL/UL Imbalance

Current cell association metrics in cellular networks only factor in the DL measurements. This leads to identical cell selections that are rather forced in the UL. The densification trend of network cells that is currently taking place, in contrast to the homogeneous settings, no longer support such a metric [50].

In the DL of HetNets, the macro coverage is dominant over the small cells due to the high transmit power disparity. The UL coverage on the other hand, is identical for all TPs since terminal power is the same for all. It is therefore easy to deduce that the naive association methods based on the Reference Signal Received Power (RSRP) will cause an imbalance in the UL/DL [51]. For HetNet tiers with less transmit power difference between the macro and small cells, the UL/DL imbalance problem is less prominent. Also, RE is shown to reduce such

effect as the positive bias value used will increase the coverage area of the macro. However, for a *coupled access*, studies in [52] [53] report a deterioration in gains as interference in the DL grows for bias values larger than 6 dB making the architecture more dependent on an appropriate ICIC technique. Furthermore, there is also an imbalance in load between the macro and the low power cells [50]. The forced UL association based on DL RSRP is therefore evidently sub-optimal. To this end, decoupling the UL/DL is motivated by both the imbalances above for improved system performance.

The above limitation on the RE will be alleviated through facilitating the decoupled access. Gains experienced through RE will continue to increase with increasing the bias values with no side effects of the DL interference any more.

Major enhancements in cell-edge UL throughput and maintaining a load-balanced system are shown as part of the main contributions of decoupling the UL/DL. Expected increase in throughput is witnessed as desired signal quality rise and the interference levels drop in such network scenarios with decoupled access. Up to Release 12, no formal reports were issued by 3GPP to draw comparisons showing the suitability of UL/DL split. UL improvements are discussed and shown in [54] for a co-channel deployment of HetNets. Further remarkable performance gains are shown in [55] for various load scenarios in the UL.

The trend in technology is driving the need for a faster and a more connected network. An integral part of the future 5^{th} Generation (5G) cellular networks and a viable solution to provide for the future "data shower" is the shift towards a centralised network infrastructure [91] with a highly evolved cellular network architecture as cloud radio access networks (C-RAN) to facilitate decoupled processing and transmission [92]. This will change the dynamics in such ultra-dense networks forming a highly heterogeneous nature. Ignorant to the uplink-downlink transmit power disparity, current downlink cell association metric designs impose significant capacity-impairing effects on uplink traffic.

On the other hand, through coordinated transmission, due to the increased levels of signal strength in the cell-edge, the needs for transmitting at higher powers can become increasingly less felt while achieving a two-fold purpose of increased capacity and reduced interference on neighbours. However, as capacity of such small cells in the HetNet environment soars, the capacity-limited fronthaul data pipes might not be able to support the naive decoupling of UL and DL in presence of CoMP. Cell association management therefore plays a paramount role specially in conjunction with CoMP's fronthaul link bottleneck limitation considered [93]. To tackle this problem in presence of CoMP, as a technology offering higher degrees of connectivity and capacity [94] further research is needed since decoupled access has only been investigated in the literature in the absence of coordinated transmission.

The downlink-uplink decoupling (DUDe) concept investigated in [95], [91], [96] and [97] for HetNets considers the power disparity amongst the evolved Node Bs eNBs contrary to the previous downlink-oriented association metrics. Based on DUDe, power-limited UL users' associations will depend on the path loss and will be decoupled from the DL association based on the received-power. Decoupling the UL and DL associations will introduce significant benefits to uplink traffic. Authors in [98] utilise the notion of DUDe to present UL gains through the decoupling of UL and DL. Mobile handover for both a noise-limited and an interference-limited scenario is simulated and studied. Supported by Vodafone's live Long Term Evolution (LTE) field trial network, UL throughput gains of up to 3 times the 5^{th} percentile for dense network configurations were shown in this work. However, no realistic upper-bound is considered in this work on traffic flow through the backhaul links. A mathematical framework is presented for a two-tier HetNet in [99] for the association and coverage probability. Authors model the HetNet using stochastic geometry and show that the decoupling probability will increase as the density of small cell eNBs rises with respect to the macro

base stations. Authors also show that higher UL/DL decoupling degrees lead to higher UL coverage probabilities. However, not considering the limit on the aggregate traffic traversing to/from the core network, the association assumed in [98] and [99] will draw traffic to the high power macro eNBs in the DL and nearby small cell eNBs in the UL for users exhausting the associated links to the core network. In [100] and [101], cell association is examined, however, UL considerations are omitted in these works. In [102], authors advance to consider cell load and backhaul capacity in the DUDe-based cell association metric making use of high resolution path loss prediction. Different power control settings have been employed in this work and a significant 10-15 dB of UL signal to interference and noise ratio (SINR) variance is shown over a baseline model. However, authors in [102] only suffice to consider an aggregate backhaul capacity constraint whereas, in a realistic network scenario, such constraints will apply to all available links. Additionally, the considered association metric will be limiting to a network supporting joint transmission (CoMP). Finally, [103] provides a strong analytical framework for the capacity gains of DUDe for single transmission (non-CoMP) systems through the employment of stochastic geometry.

Chapter 3

Joint Packet Scheduling for the Downlink of LTE

3.1 Introduction

Managing the detrimental effects of interference in the cell-edge in the cellular networks has attracted significant research attention. To this end, increasing the cell-edge user throughput in the LTE systems is of great importance for enhanced user QoE. The requirement for higher data rates specially at the cell edges is becoming ever more pressing as the advantages could help satisfy better QoS on an end-user level and mean higher profitability on the operator side. Therefore developing algorithms and techniques to mitigate the inherently increased ICI and reduced SINR in the cell-edge is a challenge with high potential benefits and rewards. To serve the above objective, I study the joint transmission and coordinated DL packet scheduling in an LTE system under PF and full frequency reuse. We present our link-adaptive PF cooperative algorithm and show cell-edge performance improvements benchmarked against baselines. To this end, I propose a link-adaptive cooperative packet scheduling algorithm in the DL of a Long Term Evolution network employing CoMP joint transmission and coordinated

scheduling. The objective is to enhance system performance specially in the cell-edge under PF scheduling. We consider full frequency reuse and an equal-power policy. The low-complexity proposed dual-eNB PS algorithm is presented and benchmarked against the non-cooperative and the optimal fully-cooperative power-adaptive exhaustive search allocation counterparts. Both system and cell-edge throughputs are considered in the above comparison.

3.1.1 LTE Frame Structure

Peak data rate requirements of 100 Mbps for DL and 50 Mbps in the UL inspire the physical layer of LTE. Based on orthogonal frequency division multiplexing (OFDM), the frame structure supports other features such as multiple channel bandwidths [64]. The time and frequency domain frame structures are briefly discussed below.

In the *time domain*, radio resources are allocated to active-flow users every TTIs with a time duration of 1 ms also referred to as a *subframe* as depicted in Figure (3.1). An LTE frame consists of 10 subframes. Every TTI is divided into two 0.5 ms time slots each one having 7 OFDM symbols with short cyclic prefix [65]. In the *frequency domain* however, system bandwidth in the frequency domain is divided into 180 KHz *sub-channels*. Every sub-channel is made up of 12 consecutive *sub-carriers*. There is an inter sub-carrier spacing of 15 KHz in each sub-channel [66].

A time-frequency radio resource of length 0.5 ms and 180 KHz of bandwidth is referred to as a *resource block*. A resource block is the minimum resource allocation granularity unit.

The rest of the chapter is organised as follows: System model in presented in 3.2 followed by Section 3.3 in which a detailed description of the algorithm, its mathematical formulation and the corresponding components are treated. Results are discussed in Section 3.5 and finally, Section 3.6 details the concluding

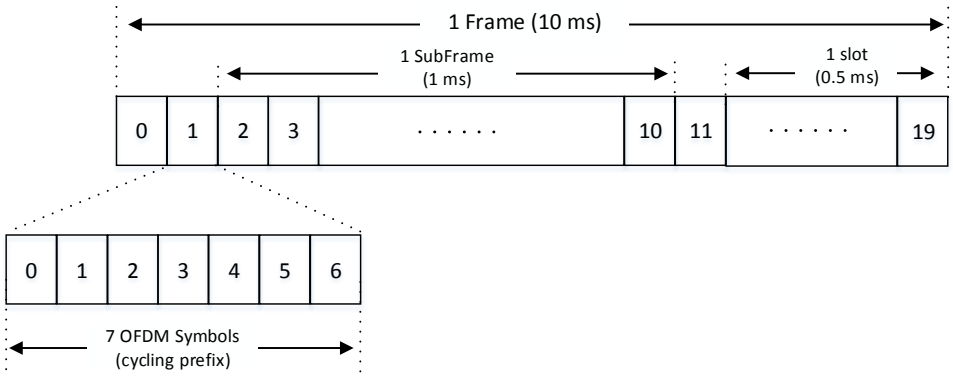

Figure 3.1: LTE frame structure.

remarks.

3.2 System Model

We consider the DL of a homogeneous multi-cell LTE system as depicted by Figure 3.2. In order to facilitate the JT and CS CoMP, I consider C-RAN as the underlying architecture for the mobile cellular network. Considering fast, low-latency and high capacity backhaul links available through C-RAN, I assume a central entity for the packet scheduling of the cells similar to [67]. The eNBs sitting at the centre of each cell perform the resource allocation and scheduling every subframe or TTI. Further, I assume perfect synchronisation between the base stations. Our considered traffic model is the infinite buffer therefore It has been assumed that there is always data available at the eNB buffer for transmission to all UEs. The sets of users, base stations and PRBs are indexed as $\mathcal{M} \triangleq \{1, ..., m, ..., M\}$, $\mathcal{N} \triangleq \{1, ..., n, ..., N\}$ and $\mathcal{K} \triangleq \{1, ..., k, ..., K\}$, respectively. Next I define two 3-D matrices for resource allocation. The time sharing matrix $\mathcal{X}_{M\times N\times K}$ with entries $x^k_{m,n} \in [0, 1]$ and the power allocation matrix $\mathcal{P} \triangleq [P^k_{m,n}]_{M\times N\times N}$. The received SINR by the m^{th} user on the k^{th} PRB at t^{th} TTI as in [C1] [1] is given by

[1] As mentioned in the Research Contributions section on page 24.

$$\gamma_m^k(t) = \frac{\sum_{\forall n \in N} |h_{m,n}^k(t)|^2 P_{m,n}^k(t) x_{m,n}^k(t)}{\sum_{\substack{\forall i \in N \\ i \neq n}} |h_{m,i}^k(t)|^2 P_{m,i}^k(t) x_{m,n}^k(t) + \delta^2}, \tag{3.1}$$

where $h_{m,n}^k(t)$, $P_{m,n}^k(t)$ and δ^2 are the instantaneous Rayleigh channel fading, transmit power and noise power over the k^{th} PRB, respectively. The time sharing factor matrix elements indicate whether PRBs have been allocated to user m by n^{th} base station on the k^{th} PRB in which case $x_{m,n}^k = 1$, and $x_{m,n}^k = 0$ otherwise. Our channel coefficient has a unit mean and considers free space path loss, shadowing and fast fading as in [57]. We also define a time set $\mathcal{T}$ to which t belongs ($t \in \mathcal{T}$). Additionally, the first term in the denominator accounts for the ICI. The instantaneously achievable rate by the m^{th} user on the k^{th} PRB as in [C1] is therefore

$$\tau_m^k(t) = B_{RB} \log_2(1 + \gamma_m^k(t)), \tag{3.2}$$

where B_{RB} is the bandwidth of a single PRB. Link adaptation or AMC adjusts the MCS used for radio links in accordance to the corresponding channel characteristics. It is vital for AMC to optimally adapt codes to radio links in order to maximise successful information transfer. Next, I present the user throughput achieved from all base stations over all time-frequency resources. Equation 3.3 gives user rate as a function of the effective SINR as

$$\tau_m(t) = \sum_{\forall n \in N} \sum_{\forall k \in K} B_{RB} \log_2(1 + \gamma_m^{eff}(t)), \tag{3.3}$$

in which, the effective SINR $\gamma_m^{eff}(t)$ takes into account the joint channel over all the PRBs utilised where

$$\gamma_m^{eff}(t) = -\beta_m^t \ln \left(\frac{1}{K} \sum_{\forall k \in K} e^{\frac{\tau_m^k(t)}{\beta_m^t}} \right), \tag{3.4}$$

Modulation	Code Rate	β
QPSK	1/3	1.49
	1/2	1.57
	2/3	1.69
	3/4	1.69
	4/5	1.65
16 QAM	1/3	3.36
	1/2	4.56
	2/3	6.42
	3/4	7.33
	4/5	7.68

Table 3.1: MCS to Beta Mapping

where β_m^t is the MCS scale factor for m^{th} user at time t [C1]. Table 3.1 lists a simple mapping between the MCS and code rate used and the beta parameter as in [68].

3.2.1 The PF Priority Function

The well-known PF scheduler's performance has been extensively examined in terms of fairness in the literature [69] [70] [71]. Here, I look at the standard PF scheduler that considers both previous and instantaneous user data rates in allocating user priorities with variable influence of the past user data rates on current priorities. Similar to [57], such a scheduler can be given a $C_{PAST}\%$ dependence to link data rates at previous TTIs for its average user rate $\bar{\tau}_m(t)$ at any TTI. For a nominal $C_{PAST}\%$, a residual $(1 - C_{PAST})\%$ will therefore be corresponding to the instantaneous data rates. The user cumulative sum of instantaneous and average past data rates for a 1:4 past-present dependence ratio (i.e. 20%-80%) is given below as in [57]

$$\bar{\tau}_m(t) = \frac{20}{100}\tau_m(t) + \frac{80}{100}\left(\frac{1}{t-1}\sum_{\substack{s \\ s\in\mathcal{T}}}^{t-1}\tau_m(s)\right). \tag{3.5}$$

As for the calculation of past average rates in the above equation, a finite window of past time intervals $[s, (t-1)]$ with length $T_{window} = t - s - 1$ can be easily defined (i.e. 10 TTIs) for the scheduler to relax the need for significant memory storage or computational complexities, suitable to the *per-TTI* nature of the packet scheduler considered. We proceed to use the above flexibility in the user priority in PF scheduling in a logarithmic [C2] [2] ratio as follows:

$$u_{ij} = Arg\max \log\left(\frac{\tau_m}{\bar{\tau}_m}\right). \tag{3.6}$$

The above scheduler will schedule UEs based on the above criterion. Equation 3.6 will be computed for all users and index of the UE maximising the above function will be returned iteratively. The log-based PF scheduler will allocate time and frequency resources to M number of UEs against the varying availability of un-utilised and unrestricted (UAU) resources in the network. The enhanced spectral efficiency of the innovated PF utility is in that, given its logarithmic nature, it will continue serving UEs with highest achievable bit rate to maximise bandwidth occupancy until a difference in orders of magnitude has been obtained in the amount of data served between the users. Only when such a trigger has been computed will the UEs with significantly lower received data take precedence over the others.

3.3 The Proposed Scheduling Algorithm

3.3.1 The Algorithm Structure

Link-adaptive DL packet scheduling of a multi-cell LTE network is investigated in this chapter. The objective of each eNB is to maximise throughput by increasing the instantaneous achievable data rate given certain constrains that will be

[2] As mentioned in the Research Contributions section on page 24.

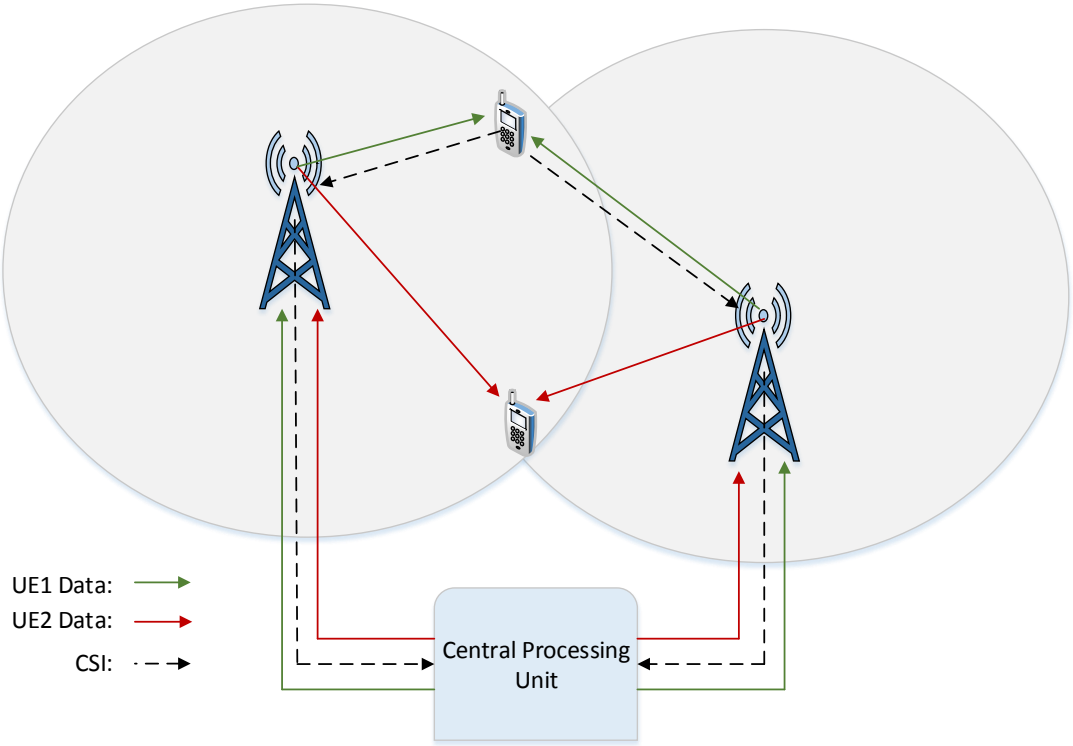

Figure 3.2: Multi-cell base station cooperation (joint JT/CS)

discussed later.

SINR computation is done at the UE end every TTI. The calculated SINRs will then be cross-checked with a quantised SINR level to map onto a 15-element long channel quality indicator (CQI) array and are then transmitted to the serving eNB. The UEs will also determine whether or not a packet is received in error based on the received SINR being above a threshold [72]. For each UE, there is correlation between the calculated SINRs over the entire bandwidth.

There are two components to the proposed scheduler: (i) the CS-CoMP component and (ii) the JT-CoMP on the physical layer. Firstly, a list of available PRBs is created and updated every TTI for all UEs. Similar to [57], based on the full-band periodic CQI reports received in each cell, eNBs map each CQI to a MCS which will also have a corresponding spectral efficiency associated with it as laid out in Table 4 of [73]. The spectral efficiency is then used to calculate the instantaneous achievable data rate using a known number of contiguous PRBs

and the number of symbols used. The utility function for all UEs over all PRBs are therefore computed and a number of PRBs from the previously updated list of available PRBs are optimally allocated to each UE based on the PF utility function considering the base station power budget and interference constraints.

3.4 Problem Formulation

In this section, I present a mathematical formulation for the scheduler proposed. The scheduler will be maximising a network utility function based on the aforementioned PF PS in Equation 3.6. Namely, the scheduling decisions are made so to maximise the sum data rate of users in the network. Limitations such as the maximum base station transmit power, minimum SINR and dual-CoMP restriction on the time sharing factor (i.e. a maximum of 2 base stations cooperating) are also noted. The notion of time (t) is dropped herein for simplicity. We present the scheduler mathematically below as the following optimisation problem [C1]:

$$
\begin{aligned}
\max_{\mathcal{X}} \quad & \mathcal{U}(\mathcal{X},\mathcal{P}) = \sum_{m\in\mathcal{M}} \log\left(\frac{\tau_m}{\bar{\tau}_m}\right) \\
\text{subject to} \quad & (I): -\beta_m \ln\left(\frac{1}{K}\sum_{\forall k\in K} e^{\frac{\tau_m}{\beta_m}}\right) \geq \gamma_m^{min}, && \forall m\in\mathcal{M} \\
& (II): \sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} x_{m,n}^k P_{m,n}^k \leq P_n^{max}, && \forall n\in\mathcal{N} \\
& (III): \sum_{n\in\mathcal{N}} x_{m,n}^k \leq 2, && \forall m\in\mathcal{M}, \forall k\in\mathcal{K} \\
& (IV): P_{m,n}^k = \frac{P_n^{max}}{\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} x_{m,n}^k}, && \forall m\in\mathcal{M}, \forall n\in\mathcal{N}, \forall k\in\mathcal{K},
\end{aligned}
\tag{3.7}
$$

where γ_m^{min} indicates minimum SINR threshold for the m^{th} UE. Moreover, the maximum transmit power associated to all base stations are assumed to be equal, that is $P_1^{max} = P_2^{max} = ... = P_n^{max}, \forall n \in \mathcal{N}$. Constraint (I) ensures individual

users' effective SINR to be above a certain threshold. Constraints (II) and (III) however, represent base station power budget and the dual-CoMP limitation. The equal transmit power policy on all PRBs considered for the algorithm is enforced by (IV). Last constraint maintains the overhead and CoMP related signalling within reasonable limits by employing a dual-eNB CoMP.

After the scheduling process, as part of the AMC to choose the best MCS for each transmission, eNBs will estimate an average SINR over the allocated PRBs from the PRB CQIs latest received from the UE based on a simple mapping. The estimated average SINR will then be used to be mapped to a MCS which will be the highest SINR-yielding scheme [74].

3.4.1 Time and Frequency Domains

Similar to [57], I consider a two-component DL scheduler at the MAC layer of cell eNB. The time domain packet scheduler (TDPS) component allocates allocation and retention priorities (ARPs) to UEs waiting to transmit based on their QCI parameters (i.e. delay tolerance) [74]. Having short-listed the UEs against UAU resources and the available resources, it then passes a request list to the frequency domain packet scheduler (FDPS). Therefore, there will always be sufficient resources to serve UEs passed down to the FDPS. Exploiting the frequency diversity [72], the FDPS will then allocate the PRBs optimally to the scheduled UEs by the TDPS considering below dynamic restrictions.

A *primary* cell is defined as the cell in which any cell-edge UE is considered for CoMP. The *secondary* cell is therefore one set to cooperate with the primary. The scheduler at the MAC layer creates a cooperation profile for all cell-edge UEs in which the primary-secondary pairs of cells are determined based on UE-eNB link quality.

If a particular cell-edge UE is short-listed for transmission at a particular TTI after the carrying out of the TDPS and FDPS, provided that it is within

the transmission range of an adjacent eNB, the eNB will be set to cooperate with the primary cell of the cell-edge UE. In addition to distance as the current parameter, other triggers can also be used to set off cooperation. In the mean time, the FDPS will have found an optimal set of PRBs for that cell-edge UE based on the latest UE-reported CQIs. Optimal time and frequency resources will be jointly chosen for cooperation between two base stations. A copy of the same packet is considered to be available at both cooperating eNBs as is the assumption in [75]. In case of the CS being triggered, the two copies of the same packet available in the two cells will be scheduled for transmission jointly at both eNBs at the same TTI to the same recipient using the same PRBs. This will significantly increase the strength of the signal at the receiver.

Joint transmission as the second component is activated only when the eNB decides to trigger joint scheduling. Upon activation, based on the ideally assumed signal profiles of the simultaneously arriving signal copies, the receiver will sum the received powers leading to a significant improvement in the received SINR. Additionally, to allow for the MRC to sum the signal powers, perfect synchronisation and channel estimation is assumed at the receiver [75]. Receiver noise is assumed to remain the same for when cooperation is applied. This simultaneous transmission using the same time-frequency resources will significantly improve the received SINR at the receiver. Also, the potential gains through our proposed algorithm will be specially obvious in applications with low delay tolerance in which the primary cell eNB will no longer have to wait for acceptable channel conditions. The cooperation advantage will allow the packet to be transmitted before it will have to be dropped due to packet delay requirements and therefore will increase throughput and decrease delay in cell-edge and satisfy the users QoS requirements.

Finally, it is important to modulate the PRBs destined to be transmitted to a particular UE using the same MCS [57].

Table 3.2: Table for Simulation Parameters

Parameter	Value
Path Loss	$32.4 + 20 \log_{10}(d_{km}) + 20 \log_{10}(f_{MHz})$
System Bandwidth	3 MHz
Carrier Frequency	2 GHz
Total Number of RBs	15
P_{eNB}^{max}	43 dBm ($\approx$ 20 W)
Shadowing Standard Deviation	8 dB
CSI	Non-varying (During a Sub-Frame) and Known
Noise Figure	2.5
Noise Spectral Density	-174 dBm
RB Bandwidth	180 KHz
Number of RB Symbols	7 (Normal Prefix)
CQI Type	Periodic and Wideband
Cell Radius	1 Km
User Distribution	Uniform and Random
TTI Duration	1 ms
T_{window}	10 ms
C_{PAST}	20%
Traffic Model	Full Buffer
Average Packet Size	500 Bytes
Simulation Duration	5000 TTIs height

3.5 Results Analysis

In this section, I present the simulated scenario parameters and simulation results for the proposed link-adaptive coordinated packet scheduler in the DL of an LTE system under the altered PF priority with full frequency reuse. Table 3.2 details a comprehensive list for the simulation parameters used.

Our algorithm's performance is compared to three other schemes. Figure 3.3 shows cell-edge spectral efficiency of the simulated scenarios versus varying maximum base station transmit power. User QoE is very much affected by the cell-edge performance. Cell-edge users that are generally more prone to interference, are scheduled on optimal resources under synchronous joint transmission of CoMP through base station coordination in our algorithm. Less power will therefore be required for reception of higher SINRs for such users. Acceptable performance loss is presented by the joint transmission and coordinated packet

scheduling under an equal-power policy in comparison to a CoMP method with adaptive power allocation in return for the significant reduction in complexity. For average base station transmit power levels of 43 dBm, there is an insignificant difference of 0.32 bits/sec/Hz in spectral efficiency as illustrated. The joint problem of optimal resource and adaptive power allocation can axiomatically question the feasibility of such algorithms that need to run on a "*per sub-frame*" basis. The superiority of the proposed algorithm over the single transmission schemes is justified as follows. As the transmit power to cell-edge users rise, both the desired signal and the interference magnitude are increased. This does not achieve a remarkable change in the SINR levels experienced in single transmission. For coordinated systems however, the desired signal power is shifted way above the interference levels as an interference source itself will now be contributing towards an elevated SINR levels as illustrated in Figure 3.4. For cooperation, the same PRBs, deemed jointly optimal are used for transmission to the same UE from two TPs. Some gains are witness in this figure for the single transmission scheme with coordinated scheduling over the plain single transmission. This is due to the incorporation of interference avoidance in the base station coordination. However, the limited SINR increase does not compare to the proposed algorithm with coordination between the TPs and the physical layer joint transmission of the packets. Dynamic PA is a powerful tool to control interference levels in the cell-edge by adjusting power levels on different PRBs hence the disparity in performance between the proposed packet scheduling and the adaptive-power counterpart.

Considering the traffic type and the average packet size considered with best effort buffers, Figure 3.5 shows the average cell-edge user packet delay for increasing number of users. For low user densities (i.e. 9), an impressing difference of 17% is shown between our proposed scheme with the dynamic power version. Inevitably, such delay is bound to increase as fixed resources are shared amongst more users. However, this increase exhibits a linear behaviour for coordinated

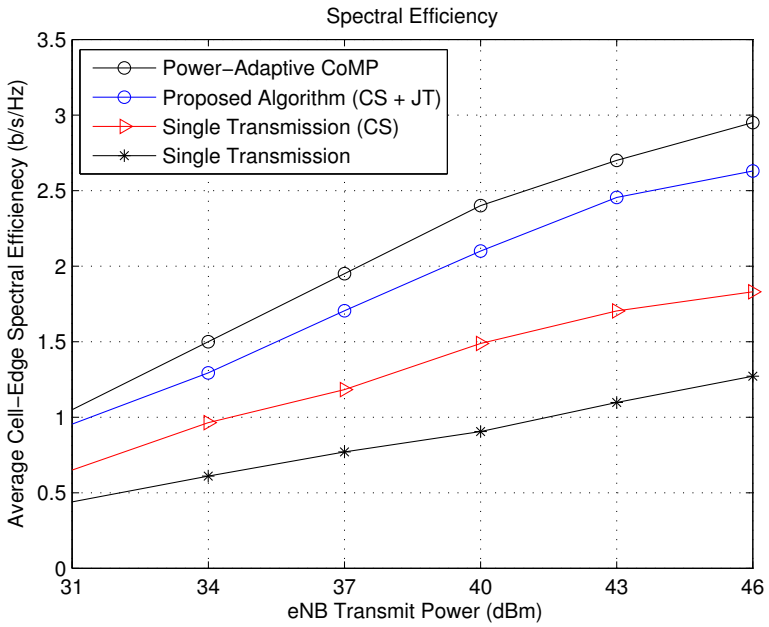

Figure 3.3: Average cell-edge spectral efficiency

transmissions as some of the performance loss in the form of buffer service rate is compensated for by the CoMP gain as more users are served.

The same performance indicator behaves rather exponentially, with various gradients for single transmission systems as the cell-edge users are deprived even more to avoid consequential drop in cell-centre user rates even under a fairness-aware user priority such as the PF. It is noteworthy to pinpoint that base station coordination does help control this adverse effect on user buffers as also shown.

Along with its gains to the primary cell-edge users, cooperation also introduces some limitations to the cooperating cells known as the secondary cells. From a signalling standpoint, feasibility of large clusters of CoMP can really be questionable due to the enormous overhead and pilot efforts required. From a radio resource management perspective however, secondary cells users will experience less frequency diversity due to the fact that a portion of the resources available to them is now utilised for joint transmission to a user residing at an adjacent cell. Delay is one of the most obvious performance factors influenced. Figure 3.6 illustrates primary/secondary cell users delay against the number of

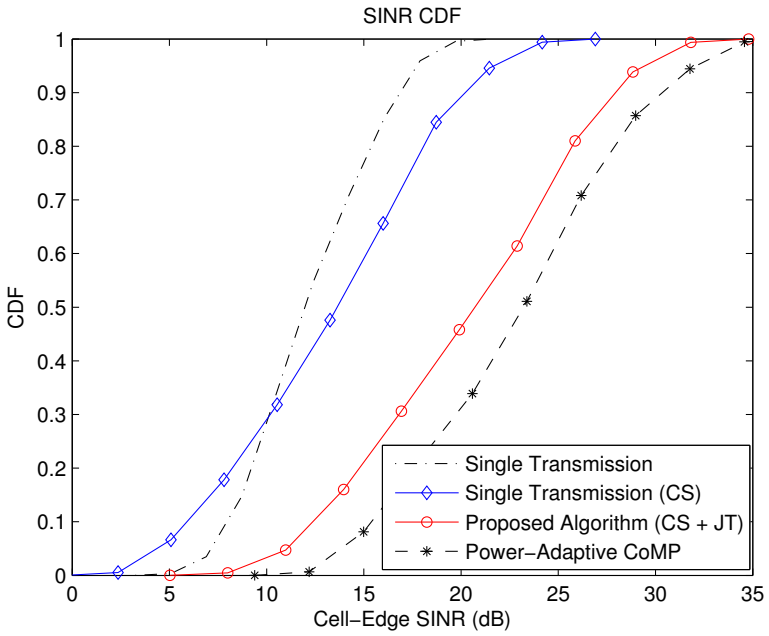

Figure 3.4: Cumulative Distribution Function (CDF)

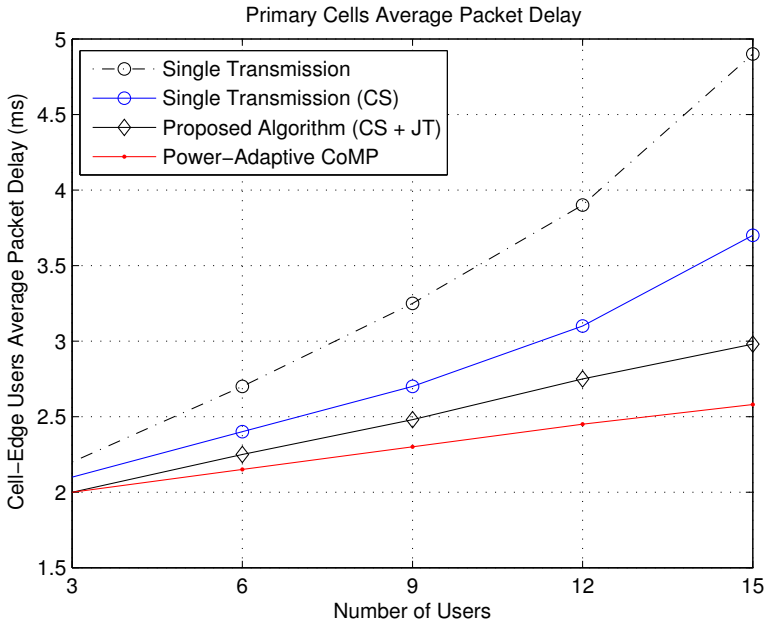

Figure 3.5: Cell-edge packet delay

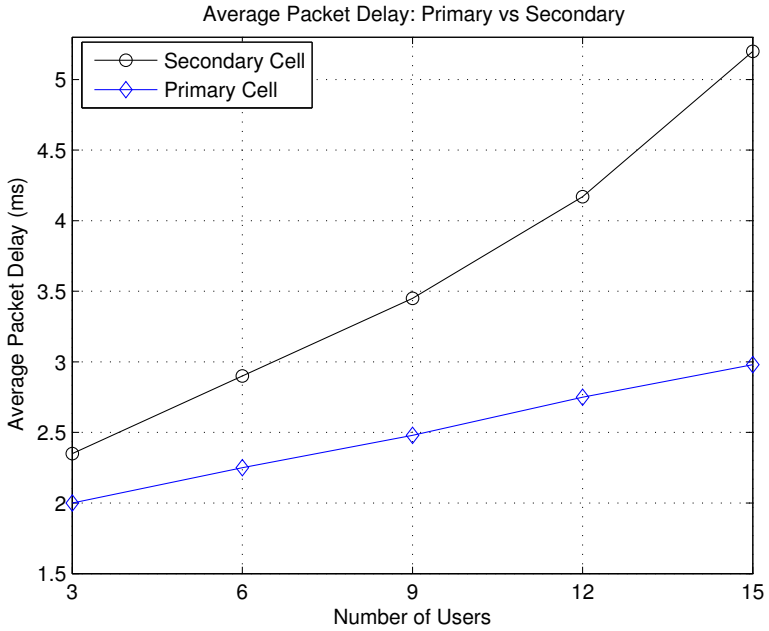

Figure 3.6: Average packet delay

users served for our proposed algorithm. Secondary users buffer as explained, will therefore experience a higher average waiting time for packet transmission due to less resources. The rationale behind the almost-linear increase in the primary cell delay follows that in Figure 3.5 as already discussed.

Figure 3.7 exhibits the average cell-edge throughput versus the number of users. On one hand, as more users compete for frequency resources, average user throughput drops. On the other hand, more cell-edge users in turn, introduce higher coordinated transmission gains translating into higher data rates, a leverage that the single transmission schemes do not enjoy. The joint effect of the two factors are characterised in this figure as mild exponential decays in which gains brought by the latter effect compensate for some of the loss in cell-edge rate due to the former. The ICI mitigation impact of adaptive power allocation is also shown by the performance gap in this figure.

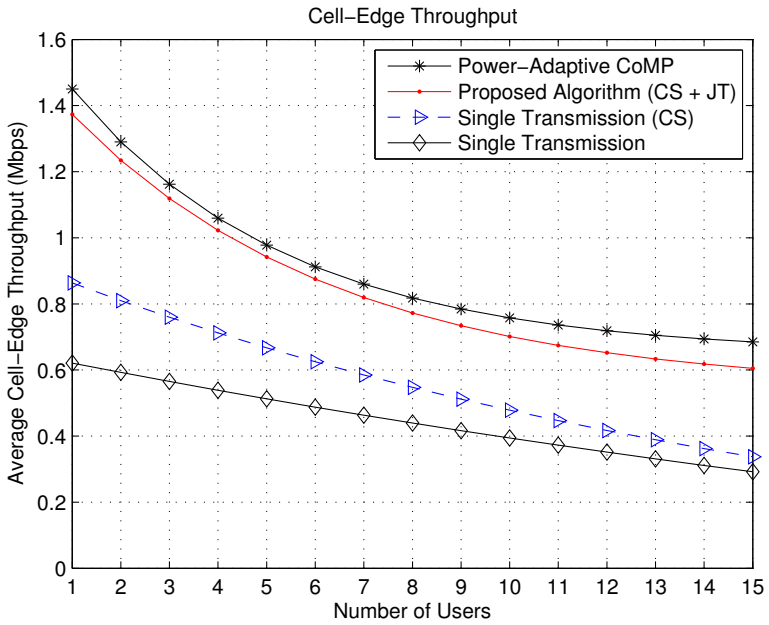

Figure 3.7: Cell-edge throughput

3.6 Conclusion

In this work, a link-adaptive coordinated packet scheduling algorithm was presented under CS and JT for CoMP. Full frequency reuse is leveraged under a PF user priority in the DL of the LTE system considered. Dynamic PA is not a consideration in the proposed algorithm in exchange for reduced complexity as a result. This facilitates the implementation of the scheme in systems whereby the scheduling decisions are made at every sub-frame, as the joint resource allocation problem with a dynamic power policy is NP-hard and not scalable. Further, a detailed structure for the algorithm was given followed by a mathematical representation formulating the objective and the scheduler constraints in resource allocation. The AMC considers user effective SINR for optimal choice of the MCS used for multiple PRBs to a user. Additionally, for the equal-power policy assumed for the scheduler proposed, our simulations illustrate reasonable reductions in gains as a result of this limitation on various performance metrics. Cell-edge enhancements inherent to CoMP methods in terms of elevated average SINR levels, higher data rates and less packet delays experienced were also

shown and analysed. The performance of our algorithm was benchmarked against the power-adaptive CoMP, coordinated but single TP transmission namely single transmission (CS) and the plain single transmission scheme. Finally, the adverse impact of base station cooperation on delay was also shown and analysed for cell users not involved.

Chapter 4

Dynamic Power Allocation and Radio Resource Management in a Multi-Carrier LTE System

4.1 Introduction

In this chapter, the aim is to study the joint problem of RRM and dynamic power allocation. Our focus will be the downlink of a homogeneous LTE environment with multi-carrier capabilities (i.e. CA). We propose a load-aware adaptive power and resource allocation (LA-APRA) algorithm for a multi-carrier system. The algorithm will consider minimum QoS provisioning and take into account the propagation characteristics of the frequency bands. Also, for improved system performance, in-band carrier load balancing is adopted. The resource allocation problem is formulated as a *MINLP* that maximizes the proportionally-fair sum rate utility function of users. The optimization problem is solved using the dual decomposition and the subgradient methods and guarantees the minimum instantaneous user rates. The numerical results exhibit significant performance gains in comparison to a number of baseline scenarios.

To this end, I study the joint problem of resource block and dynamic power allocation in the downlink of a multi-carrier LTE system. Additionally, load balancing amongst the component carriers have been shown to increase the trunking efficiency and hence a consideration of this work.

To the best of my knowledge, the joint problem of adaptive power allocation and dynamic resource management in a load-balanced multi-carrier system under minimum QoS constrains (i.e. data rate) has not been previously investigated in the literature. To this end, I propose a load-aware adaptive power and resource allocation (LA-APRA) algorithm for CC load balancing. Our strategy provides means to instantaneously guarantee the load on the CCs employed unlike Round Robin (RR) and Mobile Hashing (MH). With this objective as the main focus of this chapter, the resource management problem is formulated as the maximisation of the proportional fairness (PF) sum rate of user equipments. The standard dual decomposition and the subgradient methods are applied to solve the optimization problem.

The rest of this chapter is organised as follows: Section 4.2 outlines the system model along with some of provisional assumptions. Problem formulation is presented in Section 4.3 followed by the methodology adopted to solve the formulated problem. Results are given and analysed in Section 4.5 and finally, Section 4.6 presents the concluding remarks.

4.2 System Model

We consider a homogeneous, single cell multi-user Long Term Evolution (LTE) network. The traffic model is assumed to be full buffer as in [57]. Both users and base stations are equipped with a single antenna. The sets of users with active flows, resource blocks (RBs) and the component carriers (CCs) are indexed as $\mathcal{M} \triangleq \{1, ..., m, ..., M\}$, $\mathcal{N} \triangleq \{1, ..., n, ..., N\}$ and $\mathcal{C} \triangleq \{1, ..., c, ..., C\}$, respec-

tively. Our focus is an interference coordinated network which implies an exclusivity restriction on the RBs in RA [84]. A single RB can only be assigned to one UE at each transmission time interval (TTI) implying zero intra-cell interference. Also, I assume a null variation in the leakage of inter-cell interference through Cloud Radio Access Network (C-RAN) management inside the controller area. The inter-cell interference is therefore considered as a down-shift of the SINR. Perfect channel state information (CSI) is assumed in this work allowing for the RA task to be carried out based on the channel conditions. The SINR of the m^{th} user over n^{th} RB residing on the c^{th} CC [J1] [3] is given by

$$\gamma_{n,c}^{m} = \frac{|h_{n,c}^{m}|^2 p_{n,c}^{m}}{I_0 + \delta^2}, \tag{4.1}$$

where $h_{n,c}^{m}$ and $p_{n,c}^{m}$ are the channel fading coefficient and the transmission power on the n^{th} RB of the c^{th} CC for the m^{th} recipient UE served by j^{th} base station, respectively. The power of the additive white Gaussian noise (AWGN) is also denoted by δ^2 and the constant inter-cell interference is noted by I_0. Therefore the upper bound on the achievable instantaneous rate for m^{th} user as in [J1] is given by

$$\tau_m = \sum_{n}^{N} \sum_{c}^{C} B_{RB} \log_2(1 + \Gamma \gamma_{n,c}^{m}), \tag{4.2}$$

where B_{RB} denotes the bandwidth of a single RB and Γ is the SINR gap to account for the deviation of the achieved rate and Shannon's upper bound for rate [85]. Furthermore, obtaining the network utility function requires the computation of the PF-based priority $\frac{\tau_m}{\bar{\tau}_m}$ for every user at the beginning of each RA interval [J1] where

$$\bar{\tau}_m = \frac{1}{\bar{t}} \sum_{t-\bar{t}}^{t-1} \tau_m \tag{4.3}$$

[3] As mentioned in the Research Contributions section on page 24.

is the m^{th} user's rate at time t averaged over the $\bar{t}$ past allocations. A dual component inter-band CA is assumed in this work with 800 MHz and 2.5 GHz carrier frequencies for the 3rd Generation Partnership Project (3GPP) Rel-10 UEs to leverage CA. The most important implication of scheduling users on carriers on higher frequencies is that signals attenuate much faster at such frequencies with larger losses due to fading. As a result, such scheduling pattern is only suitable for users in near proximity of the serving base station.

4.3 Problem Formulation

In this section, I present a formal representation of the problem formulation. We define a four-dimensional power allocation cell $P \triangleq [p^m_{n,c}]_{M\times N\times C}$ in which $p^m_{n,c} \in \mathbb{R}^+$ and a resource utilization indicator cell $x \triangleq [x^m_{n,c}]_{M\times N\times C}$ where $x^m_{n,c}$ is a binary integer indicating whether a RB has been allocated. Based on the proposed algorithm, solving the dynamic power allocation problem, given the minimum rate requirements of each user implies that the adopted power allocation policy utilises the network resources to guarantee a minimum rate subject to the condition

$$\sum_m \sum_n \sum_c p^m_{n,c} \leq P_{max}, \tag{4.4}$$

where P_{max} is the power budget of base station [84] and the left hand side of the above inequality is the cumulative transmit power of the base station in the network is defined as the total transmission power on all resource blocks, component carriers and to all users.

We formulate the adaptive power, carrier and resource block allocation as a minimum QoS design with load balancing to avoid consequent reduction in capacity trunking or radio resource utilisation efficiency. On the other hand, unlike round robin (RR) load balancing [86], the task of carrier allocation is carried out combinationally to RA in our proposed algorithm ensuring higher

frequency diversity. The sum rate for network users [J1] in the DL is given by

$$\Xi(x,P)=\sum_{m}^{M}\frac{\sum_{n}^{N}\sum_{c}^{C}x_{n,c}^{m}B_{RB}\log_2(1+\Gamma\gamma_{n,c}^{m})}{\bar{\tau}_m}. \tag{4.5}$$

The objective of the LA-APRA algorithm will be to dynamically allocate power, RBs and component carriers to UEs to meet the minimum rate requirements. The joint adaptive power allocation and scheduling problem is formulated similar to my work in [J1] as a MINLP as follows:

$$\max_{x,P}\ \Xi(x,P)$$

s.t.

$$\sum_{n}\sum_{c}x_{n,c}^{m}B_{RB}\log_2(1+\Gamma\gamma_{n,c}^{m})\geq\tau_{m}^{min},\qquad\forall m\in\mathcal{M} \tag{4.6a}$$

$$1-\vartheta\leq\frac{\sum_{m}\sum_{n}x_{n,c}^{m}}{\sum_{m}\sum_{n}x_{n,v}^{m}}\leq 1+\vartheta,\qquad\forall(v,c)\in\mathcal{C},v\neq c \tag{4.6b}$$

$$\sum_{m}\sum_{n}\sum_{c}x_{n,c}^{m}p_{n,c}^{m}\leq P_{max}, \tag{4.6c}$$

$$\sum_{m}^{M}x_{n,c}^{m}\leq 1,\qquad\forall n\in\mathcal{N},\forall c\in\mathcal{C}, \tag{4.6d}$$

where $x_{n,c}^{m}\in\{0,1\}$ and $p_{n,c}^{m}\geq 0$. The constraint (4.6a) ensures the satisfaction of every UE's minimum rate condition. Constraint (4.6b) manages an instantaneous allocation of load on the CCs [86] (unlike conventional balancing schemes) in which ϑ is the load balancing slackness factor where $0\leq\vartheta\leq 1$. The case where $\vartheta=0$ is an indication of complete intolerability in load balance deviation amongst the CCs making (4.6b) tight in which case

$$\vartheta=0\iff\sum_{m}\sum_{n}x_{n,c}^{m}=\sum_{m}\sum_{n}x_{n,v}^{m}.\qquad\forall(v,c)\in\mathcal{C},v\neq c \tag{4.7}$$

On the other hand, a $\vartheta=1$ implies scheduler's complete freedom in resource allocation. The choice of φ therefore acts as a control parameter for CC fairness.

Further, power limitations are adhered to by (4.6c) and finally, (4.6d) enforces the exclusivity of the RBs in the RA within the cell. The above non-convex and combinatorial problem cannot be solved by convex optimization methods.

4.4 Methodology in the Dual Domain

We apply the Lagrangian dual decomposition method to solve the above RA problem. In order to solve the problem in continuous space, all inequality constraints are temporarily treated as equalities. The Lagrangian is formed by first relaxing the coupling constraints in the above primal problem [J1] as

$$\mathcal{L}(x,P,\lambda,\alpha,\mu,\beta,\theta) = \sum_m \sum_n \sum_c \frac{x^m_{n,c} B_{RB} \log_2(1+\frac{\Gamma|h^m_{n,c}|^2 p^m_{n,c}}{I_0+\delta^2})}{\bar{\tau}_m} + \sum_m \alpha_m \left(\Upsilon^\alpha\right) - \mu\Upsilon^\mu + \sum_{c_1}\sum_{c_2} \beta_{c_1,c_2}\Upsilon^\beta - \sum_{c_1}\sum_{c_2}\theta_{c_1,c_2}\Upsilon^\theta - \sum_n\sum_c \lambda_{n,c}\left(\sum_m x^m_{n,c} - 1\right), \quad (4.8)$$

where $(c_1, c_2) \in \mathcal{C}$ and $c_1 \neq c_2$. Additionally, for ease in notation, I define the following terms of the Lagrangian used above which will also be used later in the subgradient method [J1] in Section 4.4.2 as

$$\Upsilon^\alpha = \sum_n \sum_c x^m_{n,c} B_{RB} \log_2(1+\frac{\Gamma|h^m_{n,c}|^2 p^m_{n,c}}{I_0+\delta^2}) - \tau^{min}_m, \quad (4.9)$$

$$\Upsilon^\mu = \sum_m \sum_n \sum_c x^m_{n,c} p^m_{n,c} - P_{max}, \quad (4.10)$$

$$\Upsilon^\beta = \sum_m \sum_n x^m_{n,c_1} - (1-\vartheta)\sum_m \sum_n x^m_{n,c_2}, \quad (4.11)$$

$$\Upsilon^\theta = \sum_m \sum_n x^m_{n,c_1} - (1+\vartheta)\sum_m \sum_n x^m_{n,c_2}. \quad (4.12)$$

Lagrange dual variables are represented in matrices $\lambda \triangleq [\lambda_{n,c}]_{N\times C}$, $\beta \triangleq [\beta_{c_1,c_2}]_{C\times C}$, $\theta \triangleq [\theta_{c_1,c_2}]_{C\times C}$, vector $\alpha \triangleq [\alpha_m]_{M\times 1}$ and the constant μ. Now the

primal problem decouples into a Lagrange dual function and a dual problem. The Lagrange dual problem as in [87] is defined as

$$g(\lambda,\alpha,\mu,\beta,\theta)=\begin{cases}\max\limits_{x,P}\mathcal{L}(x,P,\lambda,\alpha,\mu,\beta,\theta)\\ s.t.\quad \forall(\lambda,\mu,\beta,\theta)\geq 0\\ \qquad \sum_m^M x_{n,c}^m\leq 1\\ \qquad x_{n,c}^m\in\{0,1\},P\in\mathbb{R}^+.\end{cases}\tag{4.13}$$

The resource utilization costs are determined by the Lagrange dual function $g(\lambda,\alpha,\mu,\beta,\theta)$ on a higher level [88] in (4.13) for given values of the Lagrangian variables. Next, the dual problem at the lower level [87] [J1] is given by

$$G=\min_{\lambda\geq 0,\mu\geq 0,\beta\geq 0,\theta\geq 0} g(\lambda,\alpha,\mu,\beta,\theta).\tag{4.14}$$

The Lagrangian in (4.8) decomposes the dual function $g(\lambda,\alpha,\mu,\beta,\theta)$ in (4.13) into $N\times C$ sub-problems that need to be solved independently, for a given set of Lagrangian multipliers $(\lambda,\alpha,\mu,\beta,\theta)$. For each resource block n residing on a component carrier c, the decomposed sub-problem is given by the following on a *per-RB* basis as in [J1]

$$\begin{aligned}\max_{x_{n,c},P_{n,c}}\mathcal{L}_{n,c}(x_{n,c},P_{n,c})=&\sum_m\frac{x_{n,c}^m B_{RB}\log_2(1+\frac{\Gamma|h_{n,c}^m|^2p_{n,c}^m}{I_0+\delta^2})}{\bar{\tau}_m}\\ &+\sum_m\alpha_m x_{n,c}^m B_{RB}\log_2(1+\frac{\Gamma|h_{n,c}^m|^2p_{n,c}^m}{I_0+\delta^2})-\mu\sum_m x_{n,c}^m p_{n,c}^m\\ &+\beta_{c_1,c_2}\left(\sum_m x_{n,c_1}^m-(1-\vartheta)\sum_m x_{n,c_2}^m\right)-\theta_{c_1,c_2}\left(\sum_m x_{n,c_1}^m-(1+\vartheta)\sum_m x_{n,c_2}^m\right)\\ &-\lambda_{n,c}\left(\sum_m x_{n,c}^m-1\right),\qquad \forall m\in\mathcal{M},\end{aligned}\tag{4.15}$$

subject to $\sum_m^M x_{n,c}^m\leq 1,\ \forall n\in\mathcal{N},\forall c\in\mathcal{C}$ where $x_{n,c}^m\in\{0,1\},\ p_{n,c}^m\in\mathbb{R}^+$. In

the above equation, $x_{n,c}$ and $P_{n,c}$ are the resource allocation and power matrices only at the specified n^{th} RB residing at c^{th} CC, respectively. In other words, $x_{n,c}$ and $P_{n,c}$ have zero entries everywhere except for the n and c entries, satisfying constraint (4.6d) [85].

4.4.0.1 Time Sharing

For non-convex problems, the solution to the dual problem serves as the upper bound of the primal problem. The difference between the true optimum and this solution is regarded as the *duality gap*. Given the integer constraints in this problem, there will be a non-zero duality gap. However, assuming *time sharing property*, as $N \times C \rightarrow \infty$, the duality gap will be asymptotically zero [89]. In our case, the time sharing condition is readily satisfied as channel fading and the noise power are assumed constant for the duration of each TTI for which the above-presented optimisation problem is solved [89].

4.4.1 Optimal Resource and Power Allocation

Furthermore, I evaluate a $\mathbf{\Lambda}_{M\times N\times C}$ function for each user (i.e. $x^m_{n,c} = 1$) in order to calculate the optimality for each of the $N \times C$ sub-problems. For each user $m \in \mathcal{M}$, I define and compute the $\mathbf{\Lambda}$ matrix with entries $\Lambda^m_{n,c}$ [J1] as

$$\Lambda^m_{n,c} = \max_{P_{n,c}} \Big[\Omega(P, \alpha, \mu, \beta, \theta)\Big], \tag{4.16}$$

subject to $p^m_{n,c} \in \mathbb{R}^+$, where Ω directly follows from (4.15) as

$$\Omega(P,\alpha,\mu,\beta,\theta) = \frac{1}{\bar{\tau}_m} B_{RB}\log_2(1+\frac{\Gamma|h^m_{n,c}|^2 p^m_{n,c}}{I_0+\delta^2}) + \alpha_m B_{RB}\log_2(1+\frac{\Gamma|h^m_{n,c}|^2 p^m_{n,c}}{I_0+\delta^2}) - \mu p^m_{n,c} + \beta_{c_1,c_2}\vartheta + \theta_{c_1,c_2}\vartheta = (\frac{1}{\bar{\tau}_m}+\alpha_m) B_{RB}\log_2(1+\frac{\Gamma|h^m_{n,c}|^2 p^m_{n,c}}{I_0+\delta^2}) - \mu p^m_{n,c} + \vartheta(\beta_{c_1,c_2}+\theta_{c_1,c_2}). \quad (4.17)$$

The index of the one user that achieves the maximum value for the $\mathbf{\Lambda}$ function over all resources $n \in N$ and $c \in C$, will receive the allocation. We therefore define our resource allocation policy as

$$x^{m^*}_{n,c} = \begin{cases} 1, & m = m^* = Arg\max_m \Lambda_{n,c}, \quad \forall n \in \mathcal{N}, \forall c \in \mathcal{C} \\ 0, & \text{otherwise,} \end{cases} \quad (4.18)$$

where m^* is the index of the user that maximises $\Lambda_{n,c}$ over all resources (i.e. optimum user allocation) and $x^{m^*}_{n,c}$ is the optimum RB allocation for the m^{th} user considering the constraints in (4.6) and the problem utility in (4.5). Equation (4.18) is computed for all users (i.e. $\forall m \in M$) and the optimum resource allocation matrix x^* is found.

Having established a method for optimal allocation of the resource blocks and the component carriers, I now adaptively adjust the transmission power on the pre-allocated resources. Our power allocation policy for m^{th} user on the pre-allocated n^{th} RB on the c^{th} component carrier as a function of the Lagrangian multipliers [J1] will be

$$p^{*m}_{n,c} = Arg\max_{P^m_{n,c}\geq 0}\Big[\Omega(P,\alpha,\mu,\beta,\theta)\Big]. \quad (4.19)$$

Further, to maximise (4.17), I differentiate the above Ω function with respect

to power $p_{n,c}^m$ [J1] yielding

$$\frac{\partial\Omega(P,\alpha,\mu,\beta,\theta)}{p_{n,c}^m} = \frac{(B_{RB}\Gamma|h_{n,c}^m|^2)(\frac{1}{\bar{\tau}_m}+\alpha_m)}{(I_0+\delta^2)\left[1+\frac{\Gamma|h_{n,c}^m|^2 p_{n,c}^m}{I_0+\delta^2}\right]\ln 2} - \mu = 0,$$

$$\therefore p_{n,c}^m = \left[\frac{B_{RB}\Gamma|h_{n,c}^m|^2(\frac{1}{\bar{\tau}_m}+\alpha_m) - \mu\delta^2\ln 2 - \mu I_0\ln 2}{\Gamma|h_{n,c}^m|^2\mu\ln 2}\right]^+ \tag{4.20}$$

$$= \left[\frac{B_{RB}(\frac{1}{\bar{\tau}_m}+\alpha_m)}{\mu\ln 2} - \frac{\delta^2 - I_0}{\Gamma|h_{n,c}^m|^2}\right]^+,$$

whereby $[.]^+$ denotes $\max(.,0)$. The transmission power allocation scheme derived in (4.20) will guarantee users minimum rate requirements [90], is evidently similar to that of a 'water-filling' policy. The iterative computation of the Lagrangian variables as a function of which p^* is found for every resource block and component carrier, will ensure the constraint (4.6a) is satisfied. This way, the optimum power allocation matrix P^* is constructed for all users over the allocated resources. Once all the sub-problems in (4.15) are solved, the Lagrange dual function $g(\lambda,\alpha,\mu,\beta,\theta)$ in (4.13) is derived by Equations (4.8) and (4.16) for the given set of $(\lambda,\alpha,\mu,\beta,\theta)$ used.

4.4.2 The Subgradient Method

To spur the values of the Lagrangian variables towards the above optima, so that the above set of equations are satisfied, I update the such values iteratively using

Algorithm 1: Load-Aware Adaptive Power and Resource Allocation (LA-APRA)

initialize $\lambda,\ \alpha,\ \mu,\ \beta,\ \theta \geq 0$
for *iter = 1 : iter*$_{max}$ **do**
  compute:
    $P \leftarrow\ (\alpha_m,\ \mu_j)$ using Equation (4.20)
    $\Lambda \leftarrow (P,\ \alpha,\ \mu,\ \beta,\ \theta)$ using Equation (4.16)
  sort Λ in ascending order $\forall(m, j)$
  construct x according to $x : Arg \max_{m} \Lambda$
  if any of $(\Upsilon^{\mu},\ \Upsilon^{\alpha},\ \Upsilon^{\beta},\ \Upsilon^{\theta}) \neq 0,\ \forall(\lambda, \alpha, \mu, \beta, \theta)$ **then**
    update Lagrangian variables for next iteration:
    $\mu_j(s+1) \leftarrow \mu_j(s),\ \beta_{c_1,c_2}(s+1) \leftarrow \beta_{c_1,c_2}(s)$
    $\alpha_m(s+1) \leftarrow \alpha_m(s),\ \theta_{c_1,c_2}(s+1) \leftarrow \theta_{c_1,c_2}(s)$
  if $\frac{\partial \mathcal{L}}{\partial \lambda} = \frac{\partial \mathcal{L}}{\partial \alpha} = \frac{\partial \mathcal{L}}{\partial \mu} = \frac{\partial \mathcal{L}}{\partial \beta} = \frac{\partial \mathcal{L}}{\partial \theta} = 0$ **then**
    update $\lambda^*, \alpha^*, \mu^*, \beta^*, \theta^*$
    return (x^*, P^*)
return (x^*, P^*)

the subgradient method similar to [J1] as follows:

$$\mu_j(s+1) = \left[\mu_j(s) - \kappa_\mu(s)\,(+\Upsilon^{\mu})\right]^{+}, \tag{4.21}$$

$$\alpha_m(s+1) = \left[\alpha_m(s) - \kappa_\alpha(s)(-\Upsilon^{\alpha})\right]^{+}, \tag{4.22}$$

$$\beta_{c_1,c_2}(s+1) = \left[\beta_{c_1,c_2}(s) - \kappa_\beta(s)(-\Upsilon^{\beta})\right]^{+}, \tag{4.23}$$

$$\theta_{c_1,c_2}(s+1) = \left[\theta_{c_1,c_2}(s) - \kappa_\theta(s)(+\Upsilon^{\theta})\right]^{+}. \tag{4.24}$$

We use the constant step sizes $\kappa_\mu, \kappa_\alpha, \kappa_\beta$ and κ_θ as discussed in [88] with the following stopping criteria

$$\Upsilon^{\alpha}=0 \iff \tau_m^{min}=\sum_n\sum_c x_{n,c}^m B_{RB}\log_2(1+\frac{\Gamma|h_{n,c}^m|^2 p_{n,c}^m}{I_0+\delta^2}), \quad \forall m \tag{4.25}$$

$$\Upsilon^{\mu}=0 \iff \sum_m\sum_n\sum_c x_{n,c}^m p_{n,c}^m = P_{max}, \tag{4.26}$$

$$\Upsilon^{\beta}=0 \iff \sum_m\sum_n x_{n,c_1}^m = (1-\vartheta)\sum_m\sum_n x_{n,c_2}^m, \quad \forall(c_1,c_2) \tag{4.27}$$

$$\Upsilon^{\theta}=0 \iff \sum_m\sum_n x_{n,c_1}^m = (1+\vartheta)\sum_m\sum_n x_{n,c_2}^m, \quad \forall(c_1,c_2), \tag{4.28}$$

for given values of the Lagrangian variables. As the focus of dual decomposition techniques are to convert the problem into one without constraints by forming a Lagrangian function derived in (4.8), parameters maximising this function will be deemed optimal as all constraints have already been incorporated within it. As for the optimality of the decision variables introduced by the primal problem, I have presented an optimal solution for the calculation of them. However, the Lagrangian multipliers that will be the roots of the unconstrained Lagrangian function are iteratively found through the subgradient method. the derivative of the Lagrange function at such optimal multipliers will be zero [J1] that is

$$\frac{\partial\mathcal{L}(x^*,P^*,\lambda^*,\alpha^*,\mu^*,\beta^*,\theta^*)}{\partial\lambda^*}=0, \tag{4.29}$$

$$\frac{\partial\mathcal{L}(x^*,P^*,\lambda^*,\alpha^*,\mu^*,\beta^*,\theta^*)}{\partial\alpha^*}=0, \tag{4.30}$$

$$\frac{\partial\mathcal{L}(x^*,P^*,\lambda^*,\alpha^*,\mu^*,\beta^*,\theta^*)}{\partial\mu^*}=0, \tag{4.31}$$

$$\frac{\partial\mathcal{L}(x^*,P^*,\lambda^*,\alpha^*,\mu^*,\beta^*,\theta^*)}{\partial\beta^*}=0, \tag{4.32}$$

$$\frac{\partial\mathcal{L}(x^*,P^*,\lambda^*,\alpha^*,\mu^*,\beta^*,\theta^*)}{\partial\theta^*}=0, \tag{4.33}$$

where $\lambda^*,\ \alpha^*,\ \mu^*,\ \beta^*$, and θ^* are the optimum Lagrangian multipliers that satisfy all constraints in the primal problem.

However, due to the integer nature of our decision variable x and non-convexities in (4.6), there will be a non-zero but reasonably small duality gap associated with the optimum solution set of the problem (i.e. $\zeta > 0$). This gap is the difference between the value of the Lagrange dual function and the dual problem as functions of the optimum Lagrangian multipliers. At convergence, the exact value of the duality gap will be equal to the difference in rate associated with a single resource allocation granularity (i.e. a single RB) under global optimum power and our proposed power policy [J1], that is

$$\zeta^{x^*P^*} = \tau(n, c), \tag{4.34}$$

where n and c are any singleton elements of $\mathcal{N}$ and $\mathcal{C}$, respectively. In the interest of finding the optimum solution set (x^*, P^*), I will compute the Lagrangian variables iteratively so as to force ζ to tend to this minimum value ($\zeta \rightarrow \zeta^{x^*P^*}$). As per our algorithm, such duality gap will be very close to zero due to the time sharing property and the analysis-based power scheme derived in (4.20) ensures optimal solution to the problem posed whereby in general, such problems are NP-hard in nature and available heuristics do not guarantee optimality in any way.

4.5 Results Presentation and Analysis

In this chapter, the downlink of a 3GPP LTE single-cell deployment is simulated for a multi carrier network. The system bandwidth is considered to be 1.4 MHz on each of the two simulated component carriers implying a total of 12 resource blocks for transmission. With the minimum distance of 60 m from each base station, 10 randomly and uniformly distributed users are considered. We assume full buffer traffic for our cross-CC packet scheduler. Our channel model consists of free space path loss, shadowing and small scale Rayleigh fading as in [57]. Our

model distinguishes between the propagation characteristics of the CCs utilised by different shadow fading losses as per Table 4.1. We find the optimal values for resource utilization and the power allocation cells (x^*, P^*) using Equations (4.18) and (4.20) by iteratively adjusting the Lagrangian multipliers to a point where the stopping criteria are satisfied as per Algorithm 1. The optimal Lagrangian variables set $(\lambda^*, \alpha^*, \mu^*, \beta^*, \theta^*)$ as the solution set to these conditions are found using the parameters (i.e. step sizes for the sub-gradient method) in Table 4.1.

In the following, I benchmark the performance of our LA-APRA algorithm under various load balancing slackness factors (i.e. ϑ) with different schemes. Firstly, I consider the baseline *fixed power resource allocation* (FP-RA) algorithm in which the transmission power is equally distributed amongst the resources regardless of the channel conditions. Secondly, *discrete power level resource allocation* (DPL-RA) considered in our comparison introduces more flexibility in power allocation by facilitating transmission on 3 power levels. And lastly, Round Robin (RR) under which the newly arrived users are scheduled on component carriers less loaded. We also present the results from a Genetic algorithm (GA).

Figure 4.1 evaluates the capability of each algorithm to meet more stringent minimum rate requirements. As the percentage of the required QoS increases starting from the 512 Kbps for each user corresponding to a 0% to 1024 Kbps (i.e. 100%), the average surplus of the user rate exceeding such QoS is illustrated. We define the rate surplus as the quota of rate exceeding the minimum rate requirements. The freedom in resource allocation under $\vartheta = 1$ yields higher degrees of frequency diversity in RA and a high consequent network utility in the form of aggregate user rate and hence greater average user surplus. Such a performance however is the result of a vague awareness of load across the CCs in contrast to the $\vartheta = 0$ in which a tight CC load is instantaneously guaranteed. The downside of such strict load balancing on the network capability to meet more stringent QoS requirements becomes more evident at high loads as in Figure 4.1.

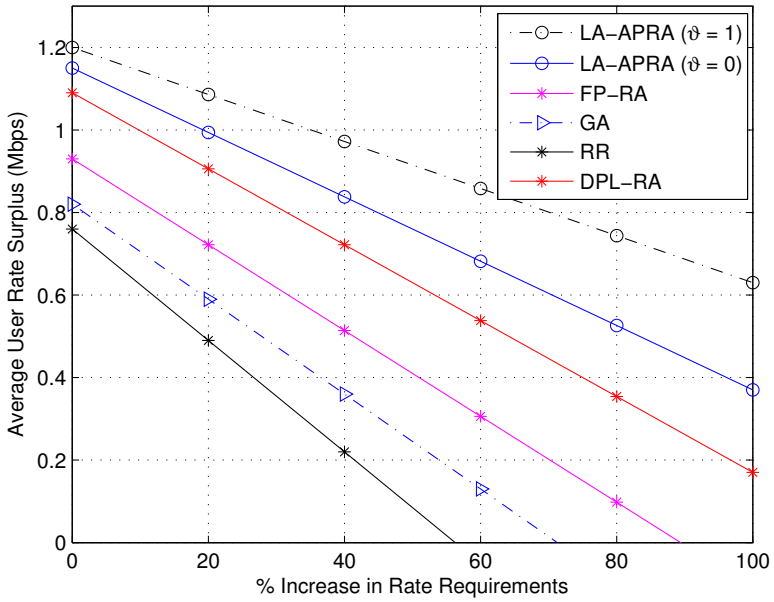

Figure 4.1: Performance under varying QoS demands for 10 UEs. User minimum bit rates (i.e. τ_m^{min}) starting from 512 Kbps.

The analytical derived expression for optimal power in our proposed LA-APRA algorithm guarantees optimality regardless of the choice of ϑ. Load balancing restrictions in effect can potentially reduce frequency diversity as a result of an optimum resource residing on a particular overly-loaded component carrier being inaccessible to the scheduler. As evident in the figure, quite predictably, DPL-RA proves more capable in providing better QoS guarantees as the optimal power levels given channel conditions will fall between two consecutive discrete power levels. With the difference between the optimal power and the chosen power level in LDPL-RA being less than that in FP-RA, this superiority in performance of DPL-RA is justified as this scheme is not as blatantly blind to channel fluctuations as the FP-RA. Finally, RR algorithm's performance scores least capable with this regard as the newly arrived users are automatically scheduled on least load CCs and this guarantees least frequency diversity for all users even in comparison to that of GA whereby an initial pre-calculated power level matrix with respect to the channel conditions and power budget is used to evaluate the user sum rate function. Such matrix is refined to maximise such function given the

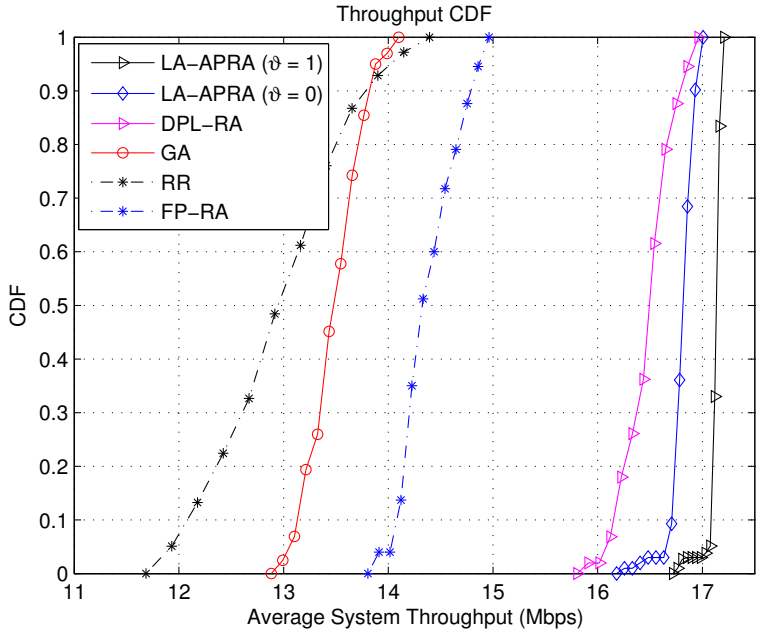

Figure 4.2: CDF for the lease (i.e. $\vartheta = 1$) and the most (i.e. $\vartheta = 0$) stringent cases with $\tau_m^{min} = 1024$ Kbps for all 10 UEs.

population size. When compared to RR, the worst case LA-APRA proves capable to guarantee more than double QoS requirements (of up to 204%) at simulated transmit power and number of users.

The cumulative distribution function (CDF) of the average system throughput is presented in Figure 4.2. The difference between the "best" and "worst" cases of LA-APRA are shown to only differ around 0.6 Mbps for a simulated scenario of 10 users with a minimum user data rate of 1024 Kbps. Best and worst terms are only used in the context of freedom in resource allocation as a result of load balancing restrictions imposed by ϑ. With similar rationale explained above for Figure 4.1, the enhance performance of the algorithm and its score against other schemes is justified as system throughput is the aggregate of the individual user rates.

As for the wider range of system throughput associated with RR, on average, a small percentage of users will not be affected by the frequency diversity limitation imposed by the RR. This can occur since such users' optimal resources will randomly happen to reside on the single CC accessible. This will consequently

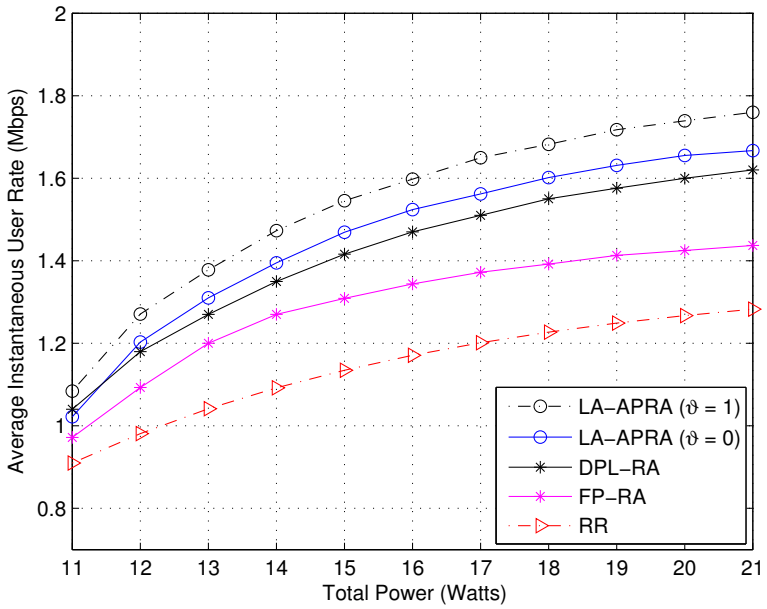

Figure 4.3: Average instantaneous rate for 10 UEs versus total transmit power with $\tau_m^{min} = 1024$ Kbps, $\forall m$.

result in a higher than average throughput well above the average users' achievable rate. However, the inverse can also occur since again a small percentage of users' performance will be affected by the RR algorithms RA mechanism more adversely than others. Shown on a CDF plot, this behaviour is demonstrated by a comparably wider interval of throughput values.

Figure 4.3 shows instantaneous user data rates under the LA-APRA algorithm against that of other baselines. The logarithmic behaviour of such graph is because the transmit power is part of the argument of the logarithmic Shannon rate function. For schemes that incorporate adjusting transmit power for changing channel gains, namely the LA-APRA and the DPL-RA, higher gradients are exhibited throughout the power ranges. Close performance between the worst case LA-APRA and the DPL-RA is due to the loss in performance of LA-APRA with $\vartheta = 0$ as a result of reduced frequency diversity being equivalent to sub-optimal power allocation performance effects under DPL-RA. Furthermore, less steep instantaneous rate increase as a function of larger transmit power profiles for FP-RA and RR are as a result of channel-unaware power allocation and

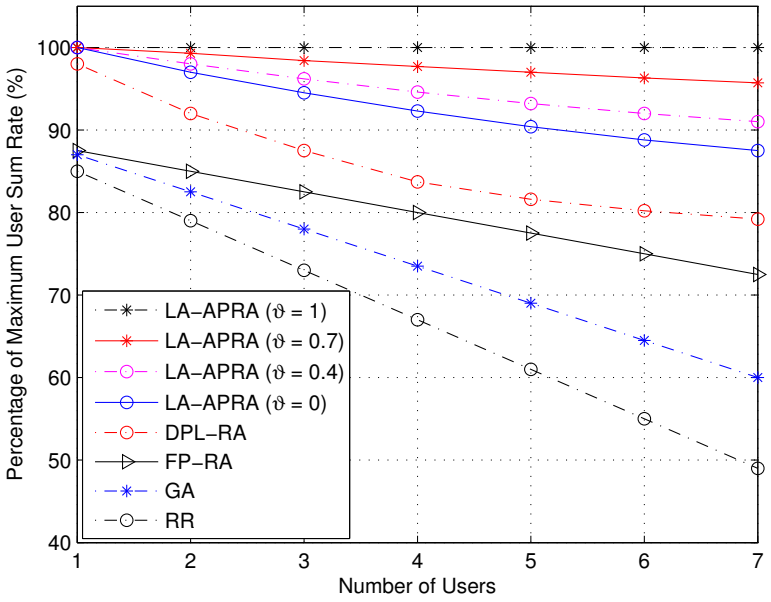

Figure 4.4: Algorithm's performance against Round Robin for varying values of ϑ for τ_m^{min} = 1024 Kbps for all users.

heavily reduced frequency diversity in resource allocation, respectively.

In Figure 4.4, I consider the loose (i.e. $\vartheta = 1$) LA-APRA as the reference scheme with respect to which other algorithms' performance in achieving sum user rate for varying number of users are benchmarked. In the single user case, the LA-APRA algorithm performs identically under all simulated values of ϑ. This behaviour is justified as the number of users scheduled on each CC is the assessment metric of load balancing and since only a single user requires scheduling, there is no distinguishing amongst different load balancing slackness factors. However, for all non LA-APRA policies, a reduced sum rate percentage, with varying drop rates, are shown for a single user scenario. As the number of users increase, the adverse effects of reduced frequency diversity associated with tighter values of ϑ on user rates become more prominent. Schemes employing tighter load balancing are susceptible to more drops in rate (i.e. drop) in comparison to the reference scheme under ($\vartheta = 1$). However, the *less-than-linear* (i.e. exponential decay) performance drops in all LA-APRA schemes and the DPL-RA algorithm for increasing users is witnessed due to the optimal power allocation

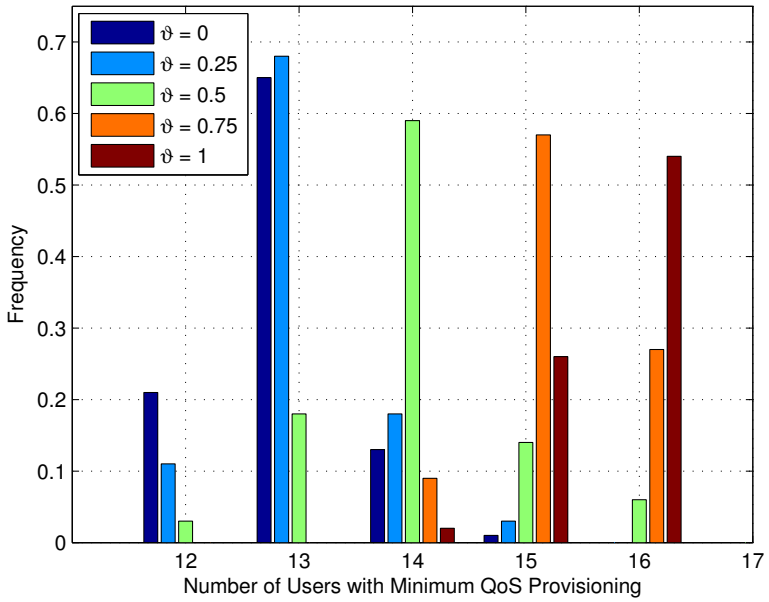

Figure 4.5: LA-APRA algorithm's capability to admit users in the network under minimum QoS ($\tau_m^{min} = 1024$ Kbps) for varying values of load balancing slackness factors (ϑ).

scheme employed, an effect that all other schemes lack. This is because some of the performance losses due to reduced higher loads and users in the network are compensated for by the sensitivity of the power allocation policies in the LA-APRA and the DPL-RA schemes. Gains associated with worst case LA-APRA in comparison to the conventional RR are verified to be up to 43% in the achieved maximum sum rate of users.

In Figure 4.5, our proposed LA-APRA algorithm's capability to meet more users with minimum QoS requirements is illustrated. The load balancing slackness factor can also be associated with *CC fairness* since the number of users are controlled through load balancing on the carriers instantaneously. Different ϑ values are compared for 500 iterations with equal user minimum data rates of 1024 Kbps. The most relaxed load balancing scheme of $\vartheta = 1$ will evidently admit largest number of users in the network with the least fairness amongst the CCs in terms of the load. For each case of ϑ, user location and distance to the base station and hence radio channel condition is the single factor determining the number of users that can be admitted in the network having their QoS re-

quirements satisfied. For the amount of bandwidth used and the transmit power budget available to the base station, an average difference of 3 users is what separates the performance of the best and worst modes of LA-APRA with other values of ϑ lying in between. This is a good indication of the performance sacrifices that can be expected for equivalent systems in order to incorporate load balancing. On average, an 18% performance gain is exhibited in worst case LA-APRA over RR in admitting more users with QoS requirements, such gains reach an average of 25% in more relaxed load balancing cases.

Table 4.1: Simulation Environment Parameters

Parameter	Value
Path Loss	$32.4 + 20\log_{10}(d_{km}) + 20\log_{10}(f_{MHz})$
Shadowing Mean	0
Shadowing Standard Deviation	8dB (@800 MHz), 10dB (@2.5 GHz)
Noise Power Spectral Density	-174 dBm/Hz
P_{max}	43 dBm ($\approx$ 20 W)
Resource Block Bandwidth (B_{RB})	180 KHz
TTI Duration	1 ms
Load Balancing Slackness Factor (ϑ)	0-1
I_0	10 dB
SINR Gap (Γ)	1
Step Sizes: $\kappa_\alpha,\ \kappa_\mu,\ \kappa_\beta,\ \kappa_\theta$	$7 \times 10^{-8},\ 25,\ 0.1,\ 0.1$
$iter_{max}$	1500
Genetic Algorithm Population Size	1200

4.6 Conclusion

In this chapter, I presented a load-aware adaptive power and resource allocation algorithm in a multi-carrier LTE system capable of instantaneous load balancing and minimum QoS guarantees. We cast the problem of joint dynamic power and resource allocation as an MINLP optimisation problem.

Given the integer nature and the non-convexities in the problem formulation, an analytical solution to the original NP-hard problem was presented using the dual decomposition and the subgradient methods. We derived a closed-form

expression for the optimum power policy that would satisfy problem constraints such as the user rate requirements, CC load and base station transmit power budget dynamically.

Benchmarked against a number of greedy heuristics, significant performance gains were exhibited in a number of performance indices. Namely, gains of up to 43% were shown in maximum sum user rates of the instantaneously load-balanced mode (i.e. worst case) of our algorithm over the conventional RR. Additionally, instantaneous user rates under our proposed scheme presented steeper growth for more generous base station power budgets. Finally, the capability of the proposed algorithm's worst case in both meeting more stringent QoS guarantees as well as admitting more QoS users in the network proved superior over the RR scheme by 204% and 18%, respectively.

Chapter 5

Fronthaul-Aware UL/DL Split Access for CoMP Systems in Heterogeneous Systems

5.1 Introduction

The general trend for the forthcoming ultra-dense 5^{th} Generation (5G) network infrastructures is directed towards higher degrees of heterogeneity. Motivated by the emergence of symmetric traffic applications, uplink is on its way to regain importance. However, in an attempt to enhance the uplink traffic flow, due to the varying transmit powers of the base stations and users in the downlink and uplink respectively, the corresponding cell associations demand a decoupled approach. Although cell association dynamics have been investigated in the recent literature for heterogeneous networks, there is no similar investigation in the presence of a capacity empowering scheme such as Coordinated MultiPoint (CoMP) with fronthaul link rate limitations. Connected to the core network through finite-capacity fronthaul links are the small cell eNBs considered in this work with different transmit power levels to the macro base stations. We present a novel

cell association algorithm for the uplink and downlink of a CoMP-enabled system and exhibit its advantages through simulations and compare results to a baseline scenario.

None of the works researched in the literature go beyond single transmission in their endeavour to approach the problem of optimal cell association. In this research, I present a novel fronthaul-aware downlink and uplink decoupling algorithm for a CoMP system based on the existing framework for DUDe. Extending the art on single transmission systems, I formulate a sum-rate maximising optimisation problem to determine optimum UL-DL associations. We also present simulations to show throughput gains brought by the decoupling of UL and UL in a CoMP scenario and illustrate how such probabilities vary for varying fronthaul rate limitations.

The rest of this chapter is organised as follows: Section 5.2 presents the system model while preliminary assumptions and the problem formulation are given in section 5.3. The proposed algorithm will be detailed in 5.4 followed by the simulations and results analysis in section 5.5. Finally, Section 5.6 will present the concluding remarks.

5.2 System Model

We consider an interference-limited heterogeneous network in LTE. The multi-cell multi-tier scenario considered is assumed to support the JT mode of CoMP by both up and downlink transmissions. Further, full buffer is employed as the traffic model in this work. It follows that in the DL, the association will be based on the highest average received power [104] from the serving base station. Uplink associations are however decoupled from those in the DL due to the disparity in transmission power between the base stations and the users. Therefore, based on DUDe, a split access will introduce significant benefits to the users in

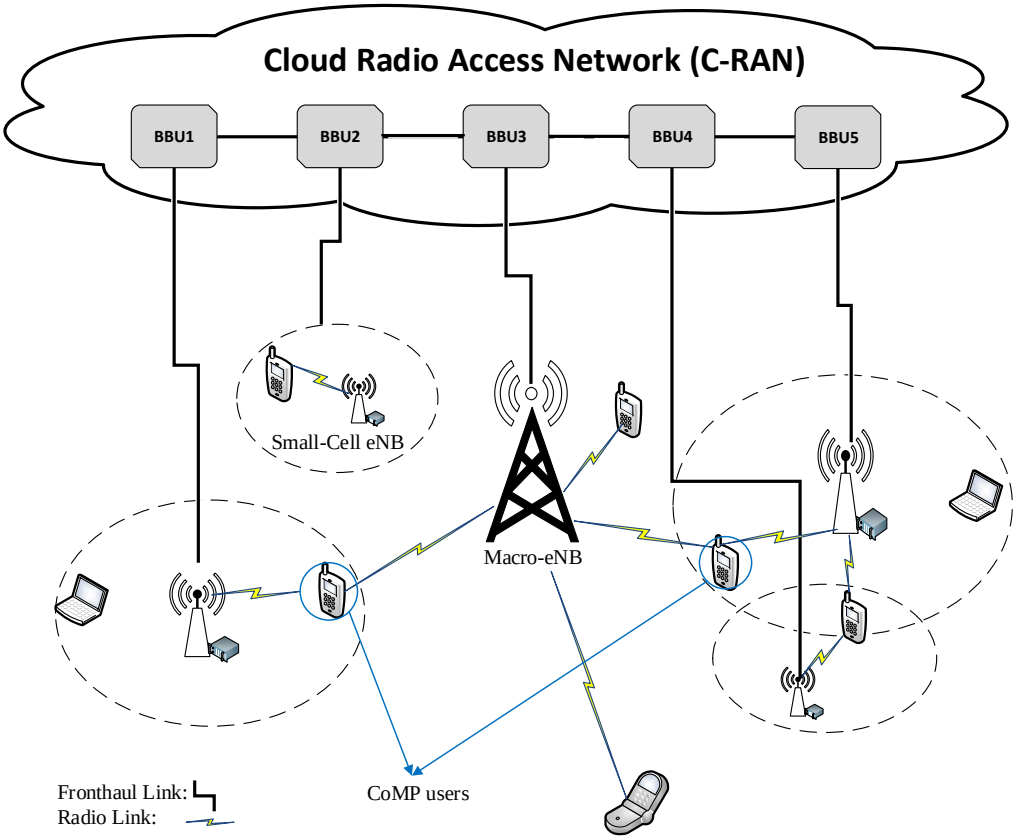

Figure 5.1: Multi-tier HetNet with fronthaul link limitation under CoMP.

the UL. On the other hand, enabling CoMP in a HetNet setting requires considerable enhancements to the current systems in the form of reduced latency amongst the cooperating TPs, tight synchronisation and the management of the overheads introduced. C-RAN as a new mobile access architecture is seen as a promising component of the future networks to facilitate such objectives. Figure 5.1 illustrates the considered scenario and the user lay-out in the heterogeneous environment and the transmission and reception of data to/from the C-RAN through finite-capacity fronthaul links. We denote the users and eNB sets by $\mathcal{K} \triangleq \{1, ..., k, ..., K\}$ and $\mathcal{J} \triangleq \{1, ..., j, ..., J\}$ respectively. We assume a separate fronthaul link associated with every eNB j ($j \in \mathcal{J}$). Similar to [103], I consider the UE-eNB received power as our association metric for UL. The CoMP-specific instantaneously received power in the UL and DL [J2] [4] are given by

[4]As mentioned in the Research Contributions section on page 24.

$$UL : W_{k,j}^{UL} = \frac{P_k^{max}}{N_k^{UL}} h_{k,j} ||d_{k,j}||^{-\alpha} + 10^{\lambda/10}, \tag{5.1a}$$

$$DL : W_{j,k}^{DL} = \frac{P_j^{max}}{N_j^{DL}} h_{k,j} ||d_{k,j}||^{-\alpha} + 10^{\lambda/10}, \tag{5.1b}$$

where $h_{k,j}$ is the instantaneous Rayleigh channel gain for the radio link between the k^{th} user and the j^{th} eNB in UL/DL and $||d_{k,j}||$, α and $10^{\lambda/10}$ denote the UE-eNB distance, path loss exponent ($\alpha > 2$) and the shadow fading ($\lambda \sim \mathcal{N}(0, \delta^2)$) terms, respectively. Also N_k^{UL} and N_j^{DL} are the total number of transmissions for the k^{th} user and j^{th} eNB, with maximum transmit powers of P_k^{max} and P_j^{max} in the UL and DL, respectively. The received power expectation will determine cell association in both DL and UL. The association criteria for the UL and DL based on the above assumptions [J2] therefore follow as

$$UL : j_k^{*UL} = arg\max \mathbb{E}[W_{k,j}^{UL}]_{j \in \mathcal{J}}, \qquad \forall k \in \mathcal{K} \tag{5.2a}$$

$$DL : j_k^{*DL} = arg\max \mathbb{E}[W_{k,j}^{UL}]_{j \in \mathcal{J}}, \qquad \forall k \in \mathcal{K}, \tag{5.2b}$$

where j_k^{*UL} and j_k^{*DL} are in turn the associated eNBs for the k^{th} UE in the UL and DL. Furthermore, notions $\mathbb{E}[W_{k,j}^{UL}]$ and $\mathbb{E}[W_{j,k}^{DL}]$ indicate the expected values for the received power in the UL and DL [J2] respectively and are given by

$$\mathbb{E}[W_{k,j}^{UL}] = \mathbb{E}[h_{k,j}] P_{k,j}^{UL} ||d_{k,j}||^{-\alpha}, \tag{5.3a}$$

$$\mathbb{E}[W_{j,k}^{DL}] = \mathbb{E}[h_{k,j}] P_{j,k}^{DL} ||d_{k,j}||^{-\alpha}, \tag{5.3b}$$

where $\mathbb{E}[h_{k,j}] = 1$. In the above, k^{th} user's transmit power on the same radio link is shown by $P_{k,j}^{UL}$ in the UL and $P_{k,j}^{DL}$ denotes the link's transmit power in the DL. In the uplink, all UEs are assumed to have equal transmit powers ($P_m^{max} = P_n^{max}$ $\forall (m, n) \in \mathcal{K}$) and therefore, the selection criterion in (5a) will only be path loss dependent. However, the received power in the downlink is a function of both

the eNB transmit power and the path loss [99].

The instantaneous SINR for the radio link between the k^{th} user and the j^{th} eNB in the UL and DL as in [J2] are given by

$$UL : \gamma_{k,j}^{UL} = \frac{\sum_j W_{k,j}^{UL}}{\sum_j \sum_{k'} W_{k',j}^{UL} + \sigma^2}, \tag{5.4a}$$

$$DL : \gamma_{j,k}^{DL} = \frac{\sum_j W_{j,k}^{DL}}{\sum_{j'} \sum_{k'} W_{j',k'}^{DL} + \sigma^2}, \tag{5.4b}$$

where σ^2 denotes the Additive White Gaussian Noise (AWGN) power.

The main impetus of this research process is the investigation of the association problem for a CoMP system. To this end, with the emphasis on the resource allocation procedure being less prominent, I assume (i) full transmit power employment for each eNB (ii) signle resource block (RB) allocation for all transmissions and (iii) equal transmit power as the power control policy on all active flows. The upper bound limit for the transmission rate of the k^{th} UE in the UL is the aggregate of rate over all receiving eNBs under the coordinated transmission scenario as assumed in [J2] and is given by

$$C_k^{UL} = \sum_{j \in \mathcal{J}} B \log_2(1 + \frac{\sum_j W_{k,j}^{UL}}{\sum_j \sum_{k'} W_{k',j}^{UL} + \sigma^2}), \tag{5.5}$$

where B is the bandwidth for a single resource block, N_k^{UL} denotes the total number of transmissions for the k^{th} user in the UL with maximum transmit power of P_k^{max}. Similarly, the aggregate instantaneously achievable rate for the k^{th} in the DL C_k^{DL} is the sum of all rates received from all eNBs [J2] and is given by

$$C_k^{DL} = \sum_{j \in \mathcal{J}} B \log_2(1 + \frac{\sum_j W_{j,k}^{DL}}{\sum_{j'} \sum_{k'} W_{j',k'}^{DL} + \sigma^2}). \tag{5.6}$$

Similarly, P_j^{max} and N_j^{DL} denote j^{th} eNB's maximum transmit power and the maximum number of radio links it transmits over in the DL, respectively.

5.3 Problem Formulation

We define, for every UE k ($k \in \mathcal{K}$) a CoMP Coalition Set (CCS) for both up (${\mathcal{S}^{\mathcal{UL}}}_k \subset \mathcal{J}$) and downlink (${\mathcal{S}^{\mathcal{DL}}}_k \subset \mathcal{J}$) transmissions to denote the set of all CoMP cooperating nodes. Such nodes will be the eNBs receiving data from the corresponding UE in the uplink and subsequently for the downlink, the eNBs jointly transmitting to the same UE. With more nodes cooperating in transmission or reception of data to/from a user, the coalition profile of that user will increase accordingly. We impose a constraint on the cardinality of the above defined CCS sets to be non-zero that is $|{\mathcal{S}^{\mathcal{UL}}}_k| > 0$ and $|{\mathcal{S}^{\mathcal{DL}}}_k| > 0$ ($\forall k \in \mathcal{K}$) indicating that each UE has at least one active flow up and downlink. The CCS cardinality will be equal to the eNB set cardinality for the case of a full-mesh connectivity that is

$$|{\mathcal{S}^{\mathcal{UL}}}_k| = |{\mathcal{S}^{\mathcal{DL}}}_k| = |\mathcal{J}| = J. \tag{5.7}$$

This represents a scenario whereby all users will be connected to all base stations in either of the DL or UL directions. Such a case is depicted in Figure (5.2). The above equation implies that all eNBs both transmit to and receive data from user k.

Next I define a user-eNB association matrix for the uplink as $\mathcal{X}^{UL} \triangleq [x^{UL}_{j,k}]_{K \times J}$ and for the downlink as $\mathcal{X}^{DL} \triangleq [x^{DL}_{j,k}]_{K \times J}$ where both $x^{UL}_{j,k}$ and $x^{DL}_{j,k} \in \{0, 1\}$. An example of the UE-eNB association matrix is given below indicating a connection pattern for 6 users and 3 eNBs

$$\mathcal{X}^{DL} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix},$$

where the matrix entry $x^{DL}_{j,k}$ in the above example will indicate a '1' to show that the radio link between the k^{th} user and the j^{th} eNB is utilised in the UL and '0' otherwise.

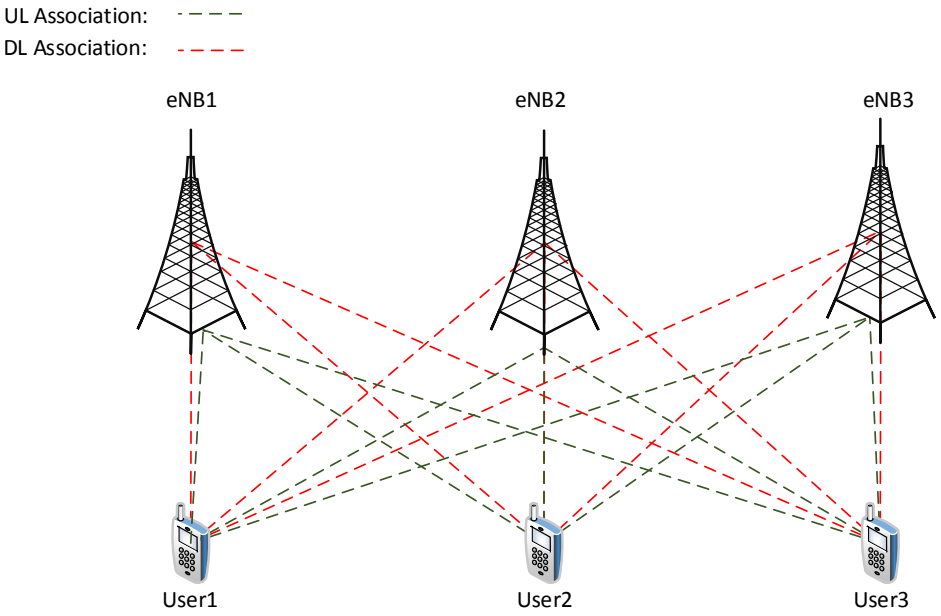

Figure 5.2: A full mesh connectivity example for both UL/DL connections.

We introduce a aggregate rate function $\mathcal{U}_k(.)$ for the k^{th} UE for both up and downlink traffic as a function of the above defined association matrices $\mathcal{X}^{\mathcal{UL}}$ and $\mathcal{X}^{\mathcal{DL}}$ similar to [J2] as follows:

$$\begin{aligned} \mathcal{U}_k(\mathcal{X}^{UL}, \mathcal{X}^{DL}) &= C_k^{UL}(\mathcal{X}^{UL}) + C_k^{DL}(\mathcal{X}^{DL}) \\ &= \sum_{j \in \mathcal{J}} x_{j,k}^{UL} B \log_2\left(1 + \frac{\sum_j W_{k,j}^{UL}}{\sum_j \sum_{k'} W_{k',j}^{UL} + \sigma^2}\right) + \\ \sum_{j \in \mathcal{J}} x_{j,k}^{DL} B \log_2\left(1 + \frac{\sum_j W_{j,k}^{DL}}{\sum_{j'} \sum_{k'} W_{j',k'}^{DL} + \sigma^2}\right), & \qquad \forall k \in \mathcal{K}. \end{aligned} \tag{5.8}$$

Given a finite amount of traffic for both up and downlink $\mathcal{L}_j$ associated with the fronthaul link of the j^{th} eNB and the above aggregate rate function, I present the following formulation for the cell association problem of a CoMP-enabled system as a linear integer optimisation problem [J2]

$$\max_{\mathcal{X}^{\mathcal{UL}},\mathcal{X}^{\mathcal{DL}}} \sum_{k\in\mathcal{K}} \mathcal{U}_k$$

s.t.

$$\sum_k x_{j,k}^{UL} B \log_2(1+\frac{\sum_j W_{k,j}^{UL}}{\sum_j \sum_{k'} W_{k',j}^{UL}+\sigma^2}) \leq \mathcal{L}_j^{UL}, \qquad \forall j \in \mathcal{J} \tag{5.9a}$$

$$\sum_k x_{j,k}^{DL} B \log_2(1+\frac{\sum_j W_{j,k}^{DL}}{\sum_{j'} \sum_{k'} W_{j',k'}^{DL}+\sigma^2}) \leq \mathcal{L}_j^{DL}, \qquad \forall j \in \mathcal{J} \tag{5.9b}$$

$$\left| \frac{\sum\limits_{k\in\mathcal{K}} C_k^{DL}(\mathcal{X}^{DL})}{C_{max}^{DL}} - \frac{\sum\limits_{k\in\mathcal{K}} C_k^{UL}(\mathcal{X}^{UL})}{C_{max}^{UL}} \right| \leq \theta \tag{5.9c}$$

$$1 \leq N_k^{UL} = \sum_{j\in\mathcal{J}} x_{j,k}^{UL} \leq J, \qquad \forall k \in \mathcal{K} \tag{5.9d}$$

$$0 \leq \sum_{k\in\mathcal{K}} x_{j,k}^{UL} \leq K, \qquad \forall j \in \mathcal{J} \tag{5.9e}$$

$$0 \leq N_j^{DL} = \sum_{k\in\mathcal{K}} x_{j,k}^{DL} \leq K, \qquad \forall j \in \mathcal{J} \tag{5.9f}$$

$$1 \leq \sum_{j\in\mathcal{J}} x_{j,k}^{DL} \leq J, \qquad \forall k \in \mathcal{K}, \tag{5.9g}$$

where constraints (5.9a) and (5.9b) ensure that each eNB's fronthaul limit for the UL ($\mathcal{L}_j^{UL}$) and DL ($\mathcal{L}_j^{DL}$) is not exceeded by maintaining the UL and DL aggregate traffic flow below these constants. The imbalance in rate between DL/UL is managed by constraint (5.9c) where the difference in ratio of each stream to its own maximum achievable capacity (C_{max}^{DL} and C_{max}^{UL}) is kept within a pre-determined bound. We define this bound as the UL/DL imbalance factor θ. Such maxima correspond to the full mesh connectivity in which all entries of UE-eNB association matrix are "1". Therefore, $\Delta^{UL/DL}$ is the imbalance associated to each unique UL/DL as per constraint (5.9d) [J2] as

$$\Delta^{UL/DL}(\mathcal{X}^{\mathcal{UL}},\mathcal{X}^{\mathcal{DL}}) = \left| \frac{\sum\limits_{k\in\mathcal{K}} C_k^{DL}(\mathcal{X}^{DL})}{C_{max}^{DL}} - \frac{\sum\limits_{k\in\mathcal{K}} C_k^{UL}(\mathcal{X}^{UL})}{C_{max}^{UL}} \right|. \tag{5.10}$$

The minimum user association constraint of at least one connection in both UL and DL is adhered to by (5.9d) and (5.9e) respectively. Finally, constraints (5.9f) and (5.9g) ensure that the maximum cell associations in the UL and DL stay within bounds. The above formulated problem will determine the optimum association matrices that maximise the aggregate of up and downlink traffic through finite capacity fronthaul links in a coordinated transmission setting. The algorithm along with a framework for the decoupling probabilities will be presented next.

5.4 Fronthaul-Aware DUDe for CoMP

5.4.1 Uplink and Downlink Decoupling

Cell association in the literature so far, has only been considered in the context of a single transmission (i.e. non-CoMP) system. We extend upon the definition for the event 'decoupled' access from the existing sources as developed for DUDe [99] [105] [95] for that of a CoMP-enabled HetNet. Such an event ($\mathcal{D}_k^S$) refers to a scenario whereby a user's uplink and downlink cell associations differ for the single transmission case that is

$$\mathcal{D}_k^S : \mathcal{S}^{\mathcal{UL}}{}_k \cap \mathcal{S}^{\mathcal{DL}}{}_k = \emptyset, \qquad \forall k \in \mathcal{K}, \tag{5.11}$$

where $|\mathcal{S}^{\mathcal{UL}}{}_k| = |\mathcal{S}^{\mathcal{UL}}{}_k| = 1$. Similarly, for any user k ($k \in \mathcal{K}$), given certain uplink and downlink CCS sets, if there exists an element in one of the sets $\mathcal{S}^{\mathcal{UL}}{}_k$ or $\mathcal{S}^{\mathcal{UL}}{}_k$ but absent in the other, such a pattern is defined as a decoupled event ($\mathcal{D}_k^C$) for this user. Therefore the CoMP counterpart of the above definition for a decoupled event is given by

$$\mathcal{D}_k^C : |(\mathcal{J} \backslash \mathcal{S}^{\mathcal{UL}}{}_k) \cap \mathcal{S}^{\mathcal{DL}}{}_k| > 0, \qquad \forall k \in \mathcal{K} \tag{5.12}$$

OR

$$\mathcal{D}_k^C : |(\mathcal{J} \backslash \mathcal{S}^{\mathcal{DL}}{}_k) \cap \mathcal{S}^{\mathcal{UL}}{}_k| > 0, \qquad \forall k \in \mathcal{K}, \tag{5.13}$$

where $\mathcal{J} \backslash \mathcal{S}^{\mathcal{UL}}{}_k$ and $\mathcal{J} \backslash \mathcal{S}^{\mathcal{DL}}{}_k$ are the complement sets of $\mathcal{S}^{\mathcal{UL}}{}_k$ and $\mathcal{S}^{\mathcal{DL}}{}_k$, respectively. We also define an optimum solution set ($\mathcal{X}^{*UL}$, $\mathcal{X}^{*DL}$) consisting of the two matrices that maximise the optimisation problem presented earlier in (5.9), given the constraints considered (i.e. fronthaul limit) and a decoupling probability $P(\mathcal{D}_k^C)$ for every user k ($k \in \mathcal{K}$) based on the above mentioned decoupling event [J2] as follows:

$$\begin{aligned} P(\mathcal{D}_k^C) &= P(\mathcal{X}_{1,k}^{*UL} \neq \mathcal{X}_{1,k}^{*DL}) + P(\mathcal{X}_{2,k}^{*UL} \neq \mathcal{X}_{2,k}^{*DL}) + ... \\ &+ P(\mathcal{X}_{j,k}^{*UL} \neq \mathcal{X}_{j,k}^{*DL}) + ... + P(\mathcal{X}_{J,k}^{*UL} \neq \mathcal{X}_{J,k}^{*DL}) \\ &= \sum_{j \in \mathcal{J}} P(\mathcal{X}_{j,k}^{*UL} \neq \mathcal{X}_{j,k}^{*DL}), \qquad \forall k \in \mathcal{K}. \end{aligned} \tag{5.14}$$

The above mathematical representation of $P(\mathcal{D}_k^C)$ implies that for a UE k to be in a decoupled state $\mathcal{D}_k^C$, the necessary and sufficient condition will be for any one of its UL/DL associations to be different. Decoupling scenarios in HetNets under both single and joint transmission schemes are illustrated in Figure (5.3). Further, I use $\mathcal{M}_{k,j}^{UL}$ and $\mathcal{M}_{k,j}^{DL}$ to denote the events of the k^{th} UE transmitting data to the j^{th} eNB (i.e. $\mathcal{X}_{j,k}^{*UL} = 1$) and receiving data from the j^{th} eNB (i.e. $\mathcal{X}_{j,k}^{*DL} = 1$) in the up and downlink [J2] respectively, with probabilities

$$P(\mathcal{M}_{k,j}^{UL}) = P(\mathcal{X}_{j,k}^{*UL} = 1) = \mathcal{P}1, \tag{5.15a}$$

$$P(\mathcal{M}_{k,j}^{DL}) = P(\mathcal{X}_{j,k}^{*DL} = 1) = \mathcal{P}2, \tag{5.15b}$$

$$\forall k \in \mathcal{K}, \forall j \in \mathcal{J}.$$

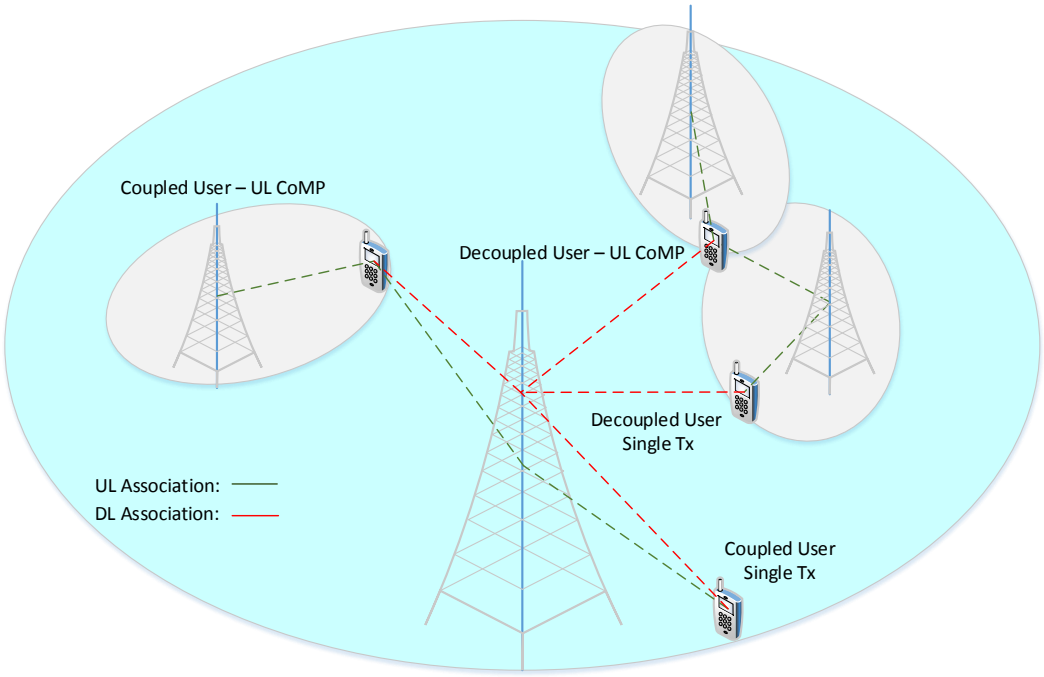

Figure 5.3: Decoupled and coupled UL/DL access instances for users under single and joint transmission cases.

Therefore Equation (5.14) becomes

$$\begin{aligned} P(\mathcal{D}_k^C) = \sum_{j\in\mathcal{J}} \Big[& P(\mathcal{X}_{j,k}^{*UL}=1)[1-P(\mathcal{X}_{j,k}^{*DL}=1)] \\ & +[1-P(\mathcal{X}_{j,k}^{*UL}=1)]P(\mathcal{X}_{j,k}^{*DL}=1)\Big] \\ = J\Big[& \mathcal{P}1+\mathcal{P}2-2\mathcal{P}1\mathcal{P}2\Big], \qquad \forall k\in\mathcal{K}, \end{aligned} \tag{5.16}$$

as in [J2].

5.4.2 The Proposed Algorithm

We propose a fronthaul-aware DL/UL decoupling algorithm for CoMP (FA-DUDe) systems that will employ any number of UE-eNB associations in both directions UL/DL to find the optimum configuration ($\mathcal{X}^{*UL}$, $\mathcal{X}^{*DL}$) for which the rate aggregate function is maximised. For every user, the algorithm will exhaust all possible association combinations as permitted by the fronthaul capacity. This

Algorithm 2: FA-DUDe for CoMP Systems

Initialise $K, J, P_k^{max}, P_j^{max}, \mathcal{L}_j^{DL}, \mathcal{L}_j^{UL}$ and iteration$_{max}$
($\forall k \in \mathcal{K}$ and $\forall j \in \mathcal{J}$)
for iteration = 1:iteration$_{max}$
Perform tasks (i) to (v):
(i): Randomly drop K UEs & J eNBs (given distributions)
(ii): Calculate cartesian distance matrices for all $(k-j)$ pair $\forall k \in \mathcal{K}$ and $\forall j \in \mathcal{J}$
(iii): Compute received power (RP) matrices according to Equation (5.1)
(iv): Compute maximum number of searchable configurations
$S_{max} = get(K, J)$
(v): Utility$_{max} = 0$
for $i = 1 : S_{max}$
(1): Compute i^{th} configuration's utility:
Utility$_i = \sum_{k \in \mathcal{K}} \mathcal{U}_k(\mathcal{X}_i^{UL}, \mathcal{X}_i^{DL})$ using Equation (8)
(2): Compute $\mathcal{L}_j^{i,DL}$ and $\mathcal{L}_j^{i,UL}$: i^{th} configuration's fronthaul utilisation $\forall j \in \mathcal{J}$ using left hand-side of constraints (5.9a) and (5.9b)
if Utility$_i >$ Utility$_{max}$
if $\mathcal{L}_j^{i,DL} < \mathcal{L}_j^{DL}$, $\mathcal{L}_j^{i,DL} < \mathcal{L}_j^{UL}$ ($\forall j \in \mathcal{J}$) and $\Delta^{UL/DL} \leq \theta$
Utility$_{max}$ = Utility$_i$
$\mathcal{X}^{*UL} = \mathcal{X}_i^{UL}$
$\mathcal{X}^{*DL} = \mathcal{X}_i^{DL}$
end if
end if
return ($\mathcal{X}^{*UL}$, $\mathcal{X}^{*DL}$)
end for
calculate and return $P(\mathcal{D}_k^C)$ $\forall k \in \mathcal{K}$
end for

will minimise the gap between the available fronthaul and the instantaneously achievable configuration rate to ensure maximum fronthaul limit utilisation and hence optimum association. Therefore, an alternative objective function to the one introduced in (5.9) can be written as

$$\min_{\mathcal{X}^{\mathcal{UL}},\mathcal{X}^{\mathcal{DL}}} \left| \mathcal{L}_j - \sum_{k\in\mathcal{K}} \mathcal{U}_k \right| \qquad \forall j \in \mathcal{J}. \tag{5.17}$$

Additionally, having ordered all achievable system capacity for feasible different configurations, our proposed algorithm guarantees an equal average and a within-bounds instantaneous imbalance in rate contribution by the UL/DL traffic depending on the value of the imbalance factor θ in constraint (5.9c) [J2], that is

$$\mathbb{E}\Big[\frac{\sum\limits_{k\in\mathcal{K}} C_k^{UL}(\mathcal{X}^{*UL})}{C_{max}^{UL}}\Big] = \mathbb{E}\Big[\frac{\sum\limits_{k\in\mathcal{K}} C_k^{DL}(\mathcal{X}^{*DL})}{C_{max}^{DL}}\Big], \tag{5.18}$$

where C_{max}^{UL} and C_{max}^{DL} denote the maximum achievable capacities associated with the *full-mesh* connectivity for the UL and DL, respectively. As with the conventional DUDe algorithm, I use the path loss for the uplink and received power for the downlink as metrics for the cell associations. As mentioned earlier, a single RB allocation with equal transmit power on each transmission is assumed. In an iterative manner, the algorithm will examine all configurations and check for optimality and based on the framework outlined in the previous subsection, the decoupling probabilities will be computed. Further, eNBs are considered to operate at maximum transmit power. Algorithm (2) details the functions of the proposed scheme.

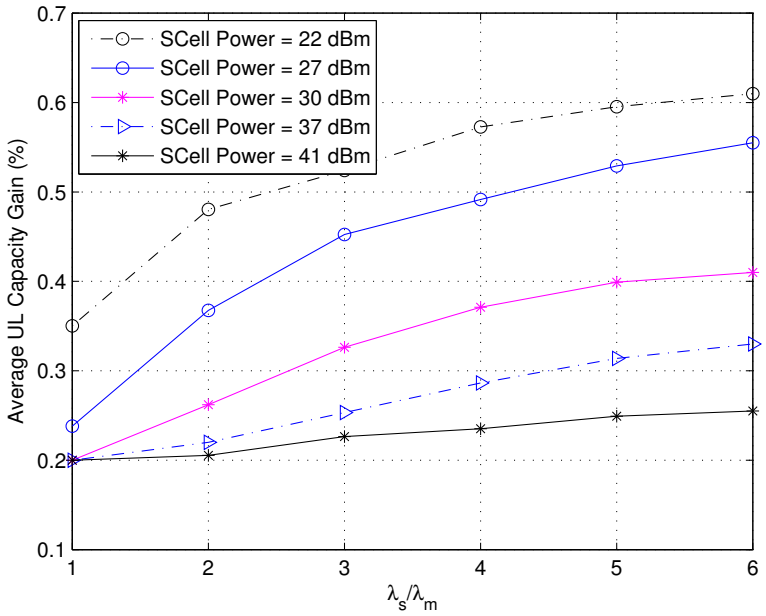

Figure 5.4: Capacity gains of the proposed algorithm (CoMP + DUDe + FH) over the coupled (CoMP + FH) counterpart for the UL users against small to macro cell density λ_s/λ_m. $K = 15$ per macro cell, $\mathcal{L}^{DL} = 1.6$ Mbps, $\mathcal{L}^{UL} = 1$ Mbps.

5.5 Simulation Environment and Results

5.5.1 Simulation Parameters

In this section, I simulate a multi-cell, interference-limited LTE HetNet consisting of users with random and eNBs with Poisson distributions [99]. The heterogeneous network considered contains a multiple tier of eNBs differentiated by their transmit powers ranging from 22 to 43 dBm corresponding to small cell and macro eNBs, respectively. Also the path loss exponent α is set to 2.5. We assume a maximum UE-eNB distance of 800m (i.e. macro cell radius). As for the bandwidth, LTE's minimum transmission bandwidth of 1.4 MHz is considered here for UL and DL. This corresponds to the usage of 6 RBs each with a bandwidth B of 180 KHz in accordance with the LTE specifications. A reasonably small value for the UL/DL imbalance factor has been chosen to be $\theta = 0.05$. Given that the total UL sum rate is significantly smaller than in the DL, large values θ will tip the balance in association towards favouring DL connections. Tight values of θ

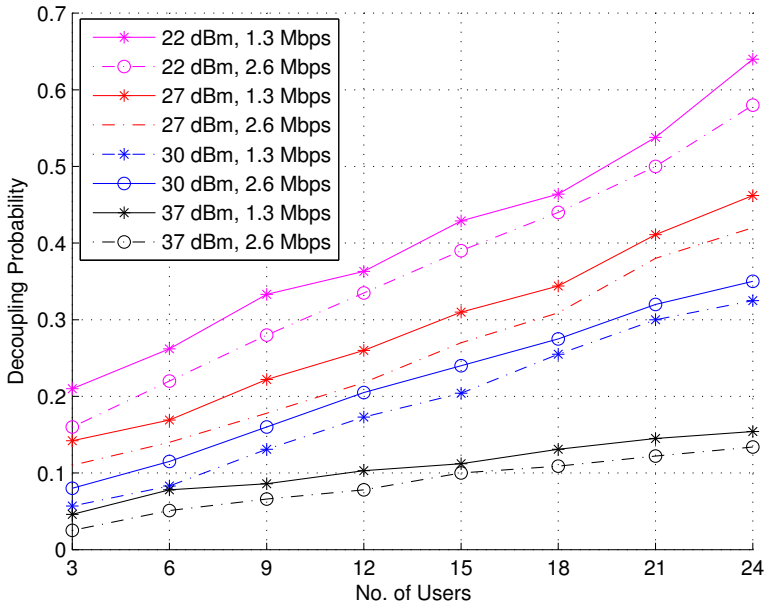

Figure 5.5: Decoupling probability of the proposed algorithm versus the number of users per macro cell for different power levels and aggregate (DL + UL) fronthaul limits. $\lambda_s/\lambda_m = 5$.

will on the other hand shrink the size of the feasible configurations and limit the gains introduced by a split access. Additionally, a shadowing standard deviation (δ) of 8 dB is used along with a noise power spectral density of -174 dBm/Hz. Next, I present the obtained simulation results conducted as per Algorithm (2) for iteration$_{max} = 10,000$.

5.5.2 Baselines Schemes

In this section, I present the baseline algorithms against which our proposed fronthaul-aware DL/UL decoupling FA-DUDe algorithm for CoMP (or *CoMP + DUDe + FH*) has been benchmarked.

CoMP will be the underlying technology for all schemes simulated here. On a broad scale, the UL and DL associations in our baselines are categorically either based upon the *downlink received power (DRP)* or are decoupled in which case they are *DUDe-based*. Further, the schemes are also distinguished from the perspective of fronthaul awareness. That is whether or not the associations

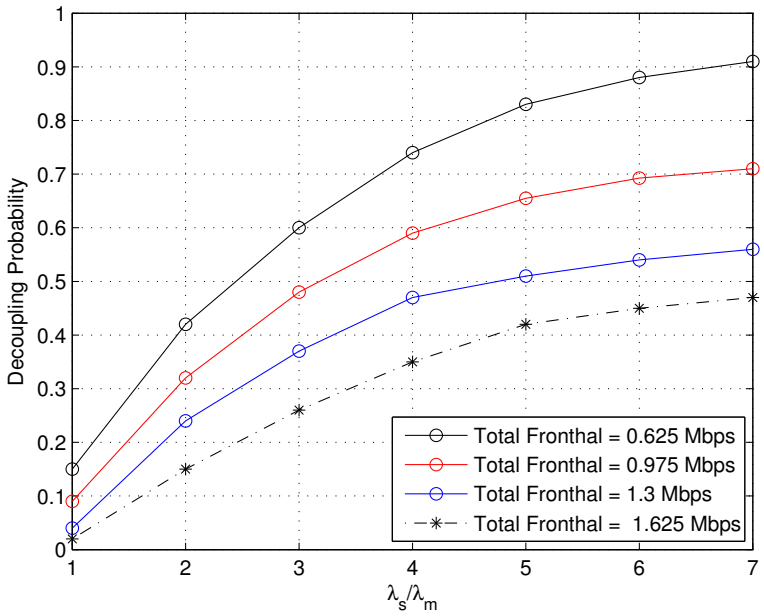

Figure 5.6: Decoupling Probability for the proposed algorithm (CoMP + DUDe + FH) versus the SCell density. $K = 20$, SCell power = 22 dBm.

between users and base stations in both uplink and downlink consider base station fronthaul capacity limitations. The simplest and the most relaxed manner of association is that based upon DRP with no fronthaul considerations (or *CoMP*). The baseline's sophistication can be further enhanced by decoupling the DL/UL, in other words, a DUDe-based CoMP (or *CoMP + DUDe*). Such an association will later be shown to achieve highest system capacity and fronthaul utilization and proves to be the *best case* association. Finally, I simulate a CoMP-based association that is based on the DRP with fronthaul considerations (or *CoMP + FH*). *Worst case* performance results will be shown to relate to the *CoMP + FH* scheme.

5.5.3 Performance Evaluation

In this section, I provide analysis to the simulations results obtained against the above-mentioned baseline scenarios.

Figure 5.4 exhibits the capacity gains introduced by the decoupling of UL/DL

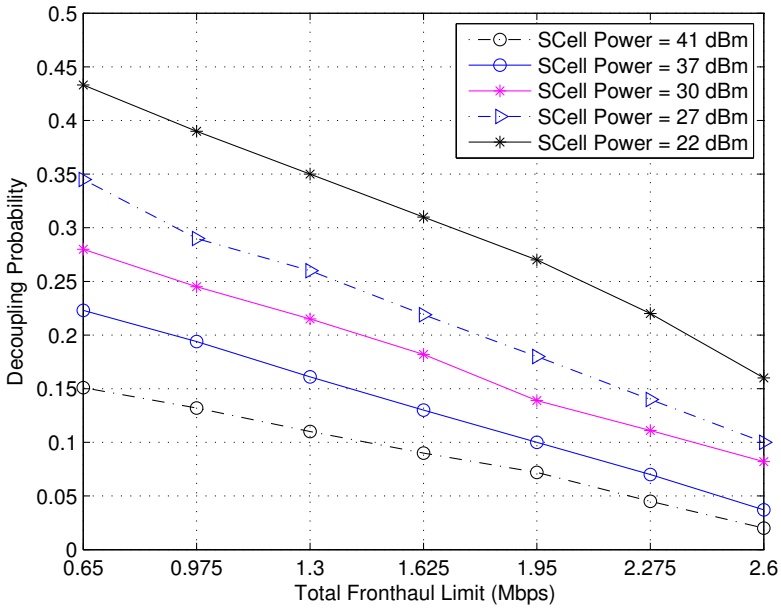

Figure 5.7: Decoupling probability of the algorithm (CoMP + DUDe + FH) against increasing fronthaul links at different SCell power levels. $K = 15$, $\lambda_s/\lambda_m = 3$.

with consideration to the capacity limitations of the fronthaul links. Over the DRP case in which UL users' associations will be dictated by those in the DL, such gains show more significant growth for higher disparities in power between the macro and the small cell base stations (i.e. SCell power). That is, for fixed macro base station power levels (i.e. as is the case in this work), lower SCell powers will achieve higher UL capacity gains. This is directly related to the high level of decoupling occurring at such low SCell power levels.

Example: Essentially, if a user is served by a base station (i.e. macro) in the DL (due to its high transmit power of 43 dBm and large coverage zone), but cannot reach the same base station in the UL (due to the user's small transmit power of 21 dBm), it will consequently choose the closest base station (i.e. with a small transmit power of 22 dBm) to connect to in the UL. This implies a decoupled access. However, if the nearest base station to this user is a base station with a larger transmit power of 38 dBm (i.e. but not a macro), then the user will be far more likely to connect to this base station in the DL than the macro base station since its signal strength can be potentially higher than that of the macro

given its closer proximity. And, the same non-macro base station with the 38 dBm power will be the user's UL association choice as well since it is the closest one. This implies a coupled access. One can deduce that probabilistically, lower SCell power levels imply larger decoupling degrees as a result of which the gains associated with lower levels of power are more prominent in the UL as shown.

This is the very motivation behind splitting access in the DL/UL as carried out by DUDe. On the other hand, for any considered power any in Figure 5.4, larger gains introduced by increasing the density ratio of small to macro cells (λ_s/λ_m) is justified as follows. With random distribution of users in the cells, the chances of users falling within the coverage areas of these small cells will naturally increase as their density increases. For a user that is not within the reachability of a such small cells, decoupled access will not be available. Therefore, higher degrees of decoupling are associated with large densities of small cells. This is one of the reasons why cell densification (i.e. large number of low power base stations) in HetNets will achieve outstanding capacity performance margins.

Decoupling probability is simulated in Figure 5.5 for varying number of users. The fact that lower SCell power levels indicate higher decoupling probabilities is justified by the same rationale explained by the above example. Additionally, more stringent fronthaul link capacities correspond to lower decoupling probabilities as shown. This is because more number of connections or radio links will have to be active in order for the fronthaul links capacities to be utilised near their maximum. This means that the system's DL and UL connection patterns will be fuller (i.e. more connections) and more similar to one another, reducing the chance of different base station associations for a given user. Finally, decoupling probability is higher for a cell with a denser user distribution since more users will be found within the coverage areas of the small cells facilitating decoupled access. The total fronthaul link capacities correspond to multiples of [0.4 - 0.25] Mbps for the $[\mathcal{L}^{DL} - \mathcal{L}^{UL}]$ pairs, in this case 2 and 4 for the cases 1.3 Mbps and

2.6 Mbps, respectively.

Figure 5.6 shows the decoupling probability of our proposed algorithm (CoMP + DUDe + FH) against varying densities of small cells. As already explained, larger densities of small cells will accommodate more users within their reachability. This will consequently allow these users to connect to such small cells in the UL shaping a different association pattern with the DL. Therefore, decoupling probability increase as a result. Furthermore, for any given density of small cell in the network, the more stringent fronthaul scenario scores higher with decoupling probability as this will not support many connections in either UL or DL. This in turn will cause a highly decoupled access as the probability that the DL/UL associations are different will be more significant.

Figure 5.7 evaluates the decoupling probability at different power levels for the proposed scheme (CoMP + DUDe + FH) versus increasing fronthaul link values. As explicitly confirmed previously in Figures 5.6 and 5.5 and also implicitly by Figure 5.4 as per the analysis, both fronthaul and SCell power levels are inversely proportional to decoupling probabilities.

The fonthaul-aware schemes' distance distributions to the serving base stations are given in Figure 5.8. Decoupled access (i.e. DUDe) ensures serving base stations will be within a closer proximity to the users. This explains the shift of the distance to the left to closer distances. The relatively narrow envelope of the is due to the fact that the coverage zone of these nodes are less than that of the macro base stations. Amongst different power levels, the shift in distance is more apparent in smaller transmit power levels as the larger power small cells (i.e. 33 dBm) are capable of associating with larger number of users that are farther away than what is witnessed for the 24 dBm case. Finally, the reason why the distance envelope is not narrower that what is illustrated is that the fronthaul awareness of the proposed algorithm (CoMP + DUDe + FH) will force some of the associations to be re-adjusted in order to accommodate the maximum fronthaul

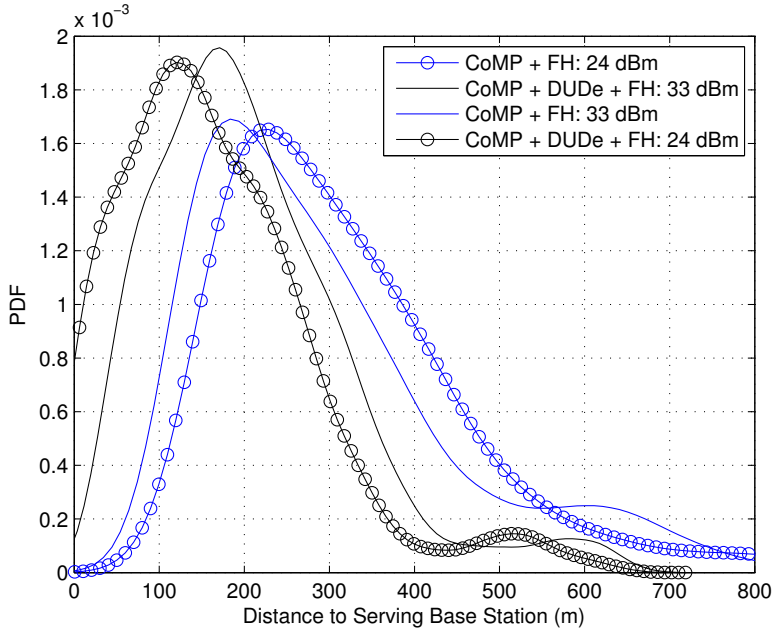

Figure 5.8: PDF for the distance to serving base stations for all four schemes. $K = 15$, $\lambda_s/\lambda_m = 5$, $\mathcal{L}^{DL} = 1.6$ Mbps, $\mathcal{L}^{UL} = 1$ Mbps

utilisation so not to violate the capacity limits. This at time, could mean that some base stations links are exhausted and nearby users will have to associate to the second best connection instead which will be further away from the user than the best connection base station.

Fronthaul utilisation of all the four schemes is illustrated in Figure 5.9 for varying levels of fronthaul capacity links. Multiples of [0.4 - 0.25] Mbps are simulated for the $[\mathcal{L}^{DL} - \mathcal{L}^{UL}]$ pairs for 2 different SCell power levels. For the plain CoMP schemes based on DRP with no fronthaul considerations, the association patterns will always be *full mesh*. This is because infinite fronthaul link capacities are assumed. Therefore, the variation of CoMP with higher power (i.e. CoMP 36 dBm) achieves more aggregate link utilisation that its lower power counterpart (i.e. CoMP 27 dBm). As for other fronthaul-unaware association policies that are DUDe-based, similar behaviour is illustrated by this figure with an increased level of link utilisation due to extra capacity brought by decoupling UL/Dl associations. A pairwise comparison (i.e. same transmit power) between DRP-based CoMP (CoMP) and the DUDe-based scheme (CoMP + DUDe) shows

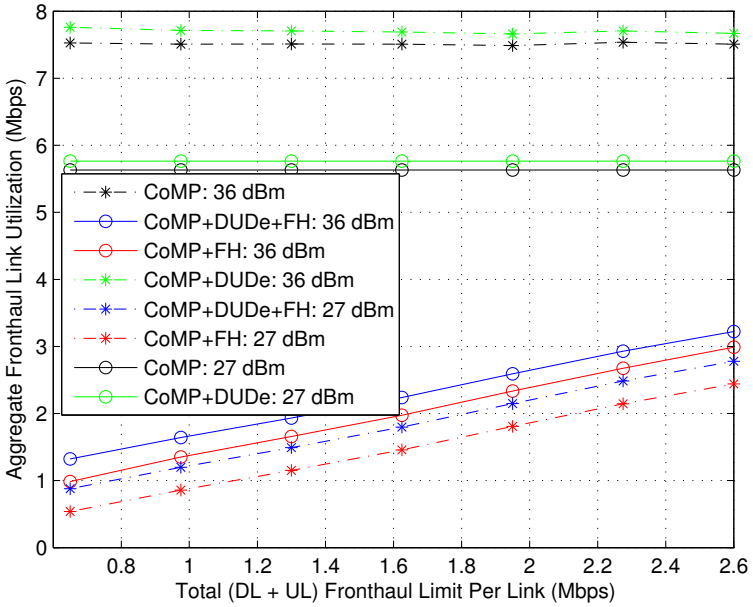

Figure 5.9: Average total fronthaul link utilisation against link limit. $K = 6$, $\lambda_s/\lambda_m = 3$.

the superiority of decoupling in obtaining higher capacity levels. Limited by the aggregate of frontahul link capacity in UL and DL, the schemes that are based on fronthaul considerations will behave differently to those outlined above. Equal power schemes of CoMP indicate more achieved capacity (i.e. link utilisation) for decoupled access as discussed previously.

Figure 5.10 exhibits the probability of joint transmission for the proposed fronthaul-aware scheme with DUDe along with the baseline algorithms. The joint transmission probability is the likelihood of a user to associate with more than one base station in UL or DL. Schemes that assume infinite fronthaul capacity will constantly maintain mesh connectivity since there will be no restrictions on the quantity of the associations they can potentially achieve. However, for fronthaul-aware schemes, such a probability will naturally increase to exhaust available link capacity as more link capacity is introduced.

Amongst such policies, our proposed scheme (CoMP + DUDe + FH) proves to rely on gains introduced by joint transmission more than the DRP-based CoMP (CoMP + FH) specially at low fronthaul limits. This is justified as follows. For

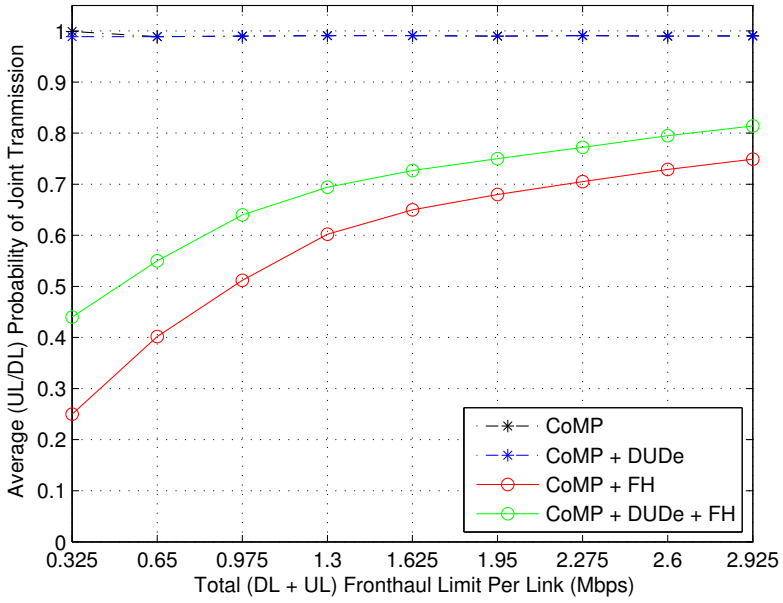

Figure 5.10: Probability of joint transmission versus different fronthaul capacity levels. SCell power = 22 dBm, $K = 30$, $\lambda_s/\lambda_m = 7$.

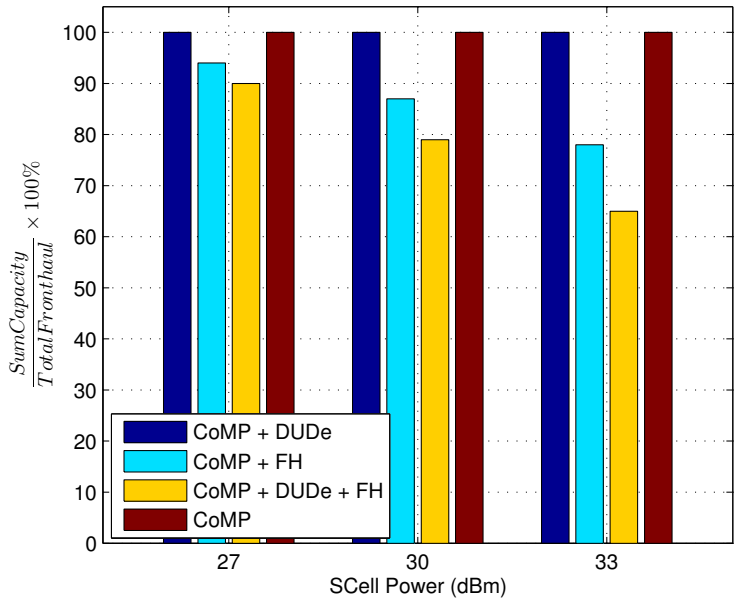

Figure 5.11: Probability of Joint Transmission - A single macro eNB with 2 small cell eNBs scenario simulated for 3 UEs.

given fronthaul limits, the DRP-based algorithm will force association in the UL to mirror that in the DL. Therefore, for an UL user to jointly transmit to multiple base stations, the DL base station, which in most cases will be a macro, will be inevitability be forced onto the UL association pattern. That is the user will need to communicate to the macro in the UL and any nearby small cells. However, for sparse cases of heterogeneity, the probability of a user to be within the coverage area of a small cell will be low, and therefore the user will only be connected to the macro base station in the UL. This is a single transmission case and the frequency of such an event is what determines the low joint transmission probability for this scheme. DUDe schemes however, have the freedom to associate with any near small cells.

Finally, the efficiency in utilising fronthaul links of the algorithms is shown in Figure 5.11. Although the fronthaul limitation is not a constraint for the fronthaul-unaware schemes (and hence the 100% efficiency throughout), for our proposed algorithm, due to the additional boost in capacity as a result of decoupled access is shown by the high efficiency in the low power levels. Compared to the higher powers, as decoupling gains drop with losing the degree of heterogeneity, efficiency shows a decrease. Such decreases are more apparent in DRP-based fronthaul-aware schemes since such gains are absent.

5.6 Conclusion

In this chapter, I investigated the problem of cell association for CoMP systems in a heterogeneous LTE environment. C-RAN was introduced as the system radio access network architecture supporting this technology given the requirement of CoMP as mentioned in Section 2.1 of the thesis specially in a high-interference scenario HetNet. The concept of decoupled access already studied in the literature under DUDe was used to establish a framework for CoMP under finite fronthaul

capacity constraints.

Further, mathematical definitions for split UL/DL access and decoupling probabilities were presented for both single and joint transmissions. We proposed our fronthaul-aware decoupling algorithm (FA-DUDe) for CoMP under which UL transmission streams would no longer suffer performance degradation due to DRP-based dictation of associations. An awareness for the UL/DL rate imbalance has been incorporated in the association strategy proposed for CoMP with the objective of maximising system sum rate. Through an operator-determined constant factor, such a imbalance is kept within reasonable bounds. An optimisation problem was formulated under which maximum users sum rate was achieved under fronthaul, association and power budget constraints. Optimal association patterns were then drawn based on the decoupled access supporting joint transmission. Additionally, baseline user/eNB association algorithms were presented and benchmarked against our proposed algorithm through extensive simulations. Various performance indicators were then illustrated to show feasibility and superiority of our proposed scheme.

Chapter 6

Conclusions and Future Work

6.1 Concluding Remarks

Maintaining acceptable user satisfaction is becoming increasing challenging for network operators. Following the unparalleled explosion in mobile data traffic, not only higher capacities are expected of wireless networks but also guaranteeing an even higher QoE in terms of connectivity and reliability for a higher volume of users. To keep pace with the projected trend, operators are resorting to the utilisation of more advanced technologies capable of satisfying such user requirements. Future generations of mobile networks will therefore be equipped with technologies that are remarkably more spectrally efficient and empower operators to meet user QoS needs. Wireless link robustness plays an integral role in cellular mobile networks. Major performance impairments are seen as a result of the existence of a amplified levels of SINR in the cell-edge. Tackling interference therefore serves as a rewarding motivation for reaching new territories in system performance. Forthcoming technologies and innovations in the communication arena are destined to traverse either of the two routes. They will have either been the direct descendent of one the prior versions through evolution and maturity or emerge as a revolutionising component into the "bloodline" of the future

networks.

Coordinated MultiPoint is a technology that addresses the problem of ICI in th cell-edge. Through base station coordination it enables converting the essence of interference sources into desired signals achieving high SINR levels and as result, enhanced system capacity. Although its introduction goes back to earlier standards, the gradual refinement of this technology in the future releases is predicted to be very promising. On the other hand, to meet the 4G requirements set by the ITU, carrier aggregation was first proposed by LTE-A to allow the use of multiple fragments of spectrum. Further, in order to fully operate at the maximum capacity possible for the wireless networks, certain architectural shifts are envisioned on a network level. Dense network deployments are proven as a seamless paradigm to guarantee a diversified range of benefits. Balancing macro load whilst covering coverage holes in the network is among the many benefits associated with them.

The prime objective and contribution of this Ph.D. thesis has been the study of the above technologies to propose new and innovative design alternatives for the purpose of achieving maximum system capacity. To accomplish this, through a background study into the existing literature, research gaps were evaluated and examined. Mathematical modelling in addition to system-level simulations were carried out so to analyse the feasibility and practicality of the design processes. Upon verification, obtained results were benchmarked against existing algorithms and strategies in favour of the hypotheses and justifications were presented and analysed. The research phases presented were conducted with the pursuit of realistic and practical assumptions in an attempt to improve currently available guidelines and schemes.

Chapter (3) investigated the problem of packet scheduling for the downlink of a CoMP system. Assuming a central scheduling entity is available for the task

of resource allocation with the full knowledge of the perfect CSI reports, link adaptation is considered in this work. Based on an effective SINR, the optimal MCS scheme on multiple RBs associated with each user is determined so to factor in the effect of varying channel gain. The MAC layer scheduler's priority is based on a PF function with weighted instantaneous and past rate considerations. The dual component equal-power packet scheduler proposed operates sequentially on both the time and frequency domains ensuring optimum allocation for active-flow users. The and the CQI are chosen as the priority basis for the time and frequency domains, respectively. Based on the CS mode of CoMP, base station cooperation is facilitated. That is a CoMP cooperation cluster is formed to decide which TPs will be cooperating. The JT mode is then activated to allow joint transmission between the determined base stations. At the receiver, MRC is employed to combine the signals arriving from different transmission nodes.

Furthermore, perfect synchronisation is assumed to be available in coordinated transmissions. Simulations are conducted based on the strategy's mathematical representation that also consider user minimum data rate requirements. User sum rate is maximised as the network utility function with equal power allocation on the resources deemed optimal by the scheduler. This reduces the algorithm complexity significantly as no additional computational arrangements are in place for dynamic allocation of power. The performance of the devised scheme for the hybrid CoMP method is compared to other main cooperation strategies. The utmost important factor for the benchmark is to illustrate the influence of optimal power policy and also base station cooperation. Two single-transmission strategies based on: 1) non-cooperative and 2) coordinated single transmission in addition to the optimal power CoMP (CS + JT) are benchmarked against our scheme. The findings of this research provide insights for performance improvements on many indicators. Whilst justification for the gains over the single transmission methods are obvious, our proposed scheme establishes an impres-

sive comparison with the dynamic (i.e. optimal) power CoMP. Considering the reduced complexity, remarkable spectral efficiency levels are obtained under our proposed scheme in comparison to its optimal power allocation policy counterpart. Substantial packet delay savings are also exhibited through cooperation based on equal power policy in comparison to its optimal power superior. Although the importance of optimal power allocation is not doubted, these findings enhance our understanding of the trade-off and the extent system performance can be compromised for reduced design complexity.

In chapter 4, with resource allocation as the main theme of this chapter, I provided an analytical basis for resource allocation in a multi-carrier system. We provided an introduction into load balancing in the context of carrier aggregation and highlighted its benefits. By comparing various existing load balancing strategies, I proceeded to establishing a mathematical representation of RRM for RB and CC allocation (i.e. non-contiguous). Whilst taking into account the need for a channel-aware power allocation, I incorporated load balancing into our analysis. Minimum QoS provisioning is also a consideration in this research process. We cast an optimisation problem with the objective of maximising a PF user sum rate and at the same time, provide an instantaneously achievable level of balance on load between the CCs (i.e. controllable). Having solved the formulated non-convex integer optimisation problem in the dual domain using Lagrangian dual decomposition and the subsequent sub-gradient methods, I derived analytical expressions for optimal power policy and the RB allocation. We also provided insights and comments on the optimality of the scheme by referencing to the time sharing property being readily available in our RA problem. The simulation findings based on our analysis of the resource allocation problem were then compared to fixed power (i.e. equal power), discrete power level, round robin and genetic algorithms. The enhanced gains through our proposed strategy were verified and further analysis was provided supporting this. Such gains were exhibited in terms

of system's capability to meet more stringent user QoS, improved instantaneous data rates and the capacity to admit a higher density of users.

Chapter (5) principally approaches the objective of capacity improvement from an optimal cell association perspective. Heterogeneous networks are critically surveyed in this chapter further as a complement to the general state-of-the-art chapter 2. With CoMP as the underlying technology, the problem of base station association was identified as a unexplored research field in the literature. The main pivotal contribution of this chapter is to shed light on the dynamics of cell association in cooperative systems. Expanding on the previously developed idea of split UL/DL access, I incorporated several valuable and realistic considerations in our research. Firstly, given that C-RAN is our assumed mobile access network to facilitate base station cooperation in a HetNet environment, I consider the fronthaul link capacity limits previously considered in major volume of research works to be non-existent.

Finally, our proposed strategy for cell association provides control on the UL/DL rate imbalance. To cater for this goal, I proceed to propose a novel cell association algorithm base upon which, cooperation between base stations are optimised so to maximise a utility function of users. The optimal association found from a list of feasible options will at the same time maintain the UL/DL rate imbalance within pre-determined bounds whilst adhering to the rate limitation of individual fronthaul links for both UL and DL. Moreover, mathematical expressions are derived for the decoupling events under both single and joint transmissions. We present our simulation results for different system parameters including heterogeneity levels (i.e. power disparity between base station tiers), fronthaul limitations, user and base station densities, etc. The performance of our fronthaul-aware method is then compared with other schemes different to each other in their incorporation of fronthaul limits and the basis for their association rules. Decoupled access for UL and DL was proposed as opposed to the associa-

tion rule based on DL received power for some of the baselines simulated. Given the tight and limiting constraints that our association strategy abides by, outstanding performance improvements are shown on UL capacity, increased SINR levels through the reduction in UE-eNB distances. The impact of decoupling was also investigated and analysed for various changing scenarios.

6.2 Future Work and Challenges

The RRM aspect of capacity enhancement in this dissertation was crucially investigated in chapters 3 and 4. In both chapters, the system models for CoMP and CA respectively, prescribed treatments for single-antenna base stations. Given the impact of interference on system throughput, it would be interesting to investigate how gains brought by CoMP and CA (individually and combined) could be capitalised by devising a scheduling strategy for the DL of a MIMO CoMP system. To tackle interference more efficiently, it would be interesting to analyse the dynamics of directional transmission facilitated through beamforming. This way, interference to neighbouring TPs is kept within bounds whilst signal strength is amplified through JT of CoMP. In addition to this, extending this work by the inclusion of the limited fronthaul link capacity concept as in chapter 5 would introduce new and exciting design prospects to the line of future research.

Resource allocation with carrier load balancing studied in chapter 4 has been limited to a single cell scenario. Although assuming a constant leakage of interference (as a down-shift in SINR) is justified through the utilisation of C-RAN, prospects of an analytical methodology for a multi-cell interference-limited system can be promising. Finally, distributed versions of scheduling algorithms for resource allocation are more attractive due to their scalability.

In the context of cell association in HetNets, the challenge of frequent HOs for a mobility scenario needs to be addressed by establishing strategies to optimally

define macro/SCell boundaries. Moreover, innovative and flexible methods are required to further suppress the impact of interference in such dense environments.

The implementation of HetNets in future releases will require handling a number of issues associated with the existing operation of current deployments.

In HetNets, HO boundaries are away from the macro cells and more toward the small low power nodes. In high mobility scenarios, terminals moving across small cells would require to acquire new cell selections frequently to maintain service. To address this, macro base station can reserve certain resources for the provision of coverage to high-mobility users. However, the interference levels from the macro coverage will overshadow small cell users. Although resource partitioning can help avoid interference introduced in this dynamic, it will also diminish frequency diversity gains brought forward by the small cell terminals. Also, fuelled by the high power disparity, UL/DL boundaries are different imposing yet another mismatch problem for the HetNets. Therefore, more emphasis needs to be focused on determining optimal HO boundaries to increase the potential of performance enhancements.

Further, solid solutions are required for the rectification of interference on the acquisition channels. Base stations with different transmit power levels that require partitioning, cause collision on the acquisition channels in the TDD mode [34]. On the contrary, the cell-specific reference signal (CRS) interference will still need tackling as the resource partitioning will only eliminate data channel interference. Mitigating CRS interference is therefore needed for both FDD and TDD modes. The performance of turbo codes and consequently the measured SINR and ultimately the gains of RE will be directly influenced by CRS interference. User-assisted CRS and acquisition signal cancellation and subtraction is also an attractive approach for managing interference for a more feasible deployment of HetNets [32].

Users in the cell-edge of the small-power nodes are most prone to macro interference. Questionable robustness in the transmission of data and control signalling to the users in the extended small cell coverage regions is another area that will need to attract some research focus in the future for more gains due to cell splitting to be witnessed [30].

Bibliography

[1] ITU-R, "Requirements Related to Technical Performance for IMT-Advanced Radio Interface(s)," ITU-R, technical report, November 2008.

[2] R1-072261, "LTE Performance Evaluation - Uplink Summary," 3GPP TSG-RAN 50, technical report, May 2007.

[3] C. Kosta, B. Hunt, A. Quddus, and R. Tafazolli, "On Interference Avoidance Through Inter-Cell Interference Coordination (ICIC) Based on OFDMA Mobile Systems," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 3, pp. 973–995, Third 2013.

[4] E. Pateromichelakis, M. Shariat, A. ul Quddus, and R. Tafazolli, "On the Evolution of Multi-Cell Scheduling in 3GPP LTE/LTE-A," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 2, pp. 701–717, Second 2013.

[5] A. Hamza, S. Khalifa, H. Hamza, and K. Elsayed, "A Survey on Inter-Cell Interference Coordination Techniques in OFDMA-Based Cellular Networks," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 4, pp. 1642–1670, Fourth 2013.

[6] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated Multipoint: Concepts, Performance and Field Trial Results," *Communications Magazine, IEEE*, vol. 49, no. 2, pp. 102–111, February 2011.

[7] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The Role of Small Cells, Coordinated Multipoint and Massive MIMO in 5G," *Communications Magazine, IEEE*, vol. 52, no. 5, pp. 44–51, May 2014.

[8] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated Multipoint Transmission and Reception in LTE-Advanced: Deployment Scenarios and Operational Challenges," *Communications Magazine, IEEE*, vol. 50, no. 2, pp. 148–155, February 2012.

[9] J. Lee, Y. Kim, H. Lee, B. L. Ng, D. Mazzarese, J. Liu, W. Xiao, and Y. Zhou, "Coordinated Multipoint Transmission and Reception in LTE-Advanced Systems," *Communications Magazine, IEEE*, vol. 50, no. 11, pp. 44–50, November 2012.

[10] M. Karakayali, G. Foschini, and R. Valenzuela, "Network Coordination for Spectrally Efficient Communications in Cellular Systems," *Wireless Communications, IEEE*, vol. 13, no. 4, pp. 56–61, Aug 2006.

[11] F. Rashid-Farrokhi, L. Tassiulas, and K. Liu, "Joint Optimal Power Control and Beamforming in Wireless Networks Using Antenna Arrays," *Communications, IEEE Transactions on*, vol. 46, no. 10, pp. 1313–1324, Oct 1998.

[12] S. Das, H. Viswanathan, and G. Rittenhouse, "Dynamic Load Balancing Through Coordinated Scheduling in Packet Data Systems," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1, March 2003, pp. 786–796 vol.1.

[13] A. Checko, H. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. Berger, and L. Dittmann, "Cloud RAN for Mobile Networks x2014;A Technology Overview," *Communications Surveys Tutorials, IEEE*, vol. 17, no. 1, pp. 405–426, Firstquarter 2015.

[14] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud Radio Access Network (C-RAN): a Primer," *Network, IEEE*, vol. 29, no. 1, pp. 35–41, Jan 2015.

[15] H. Taoka, "Views on 5G," *DoCoMo, WWRF21, Dusseldorf, Germany, Tech. Rep*, 2011.

[16] U. Rapport, "2133. Requirements, Evaluation Criteria and Submission Templates for the Development of IMT-Advanced," Technical report, January, Tech. Rep., 2008.

[17] Z. Shen, A. Papasakellariou, J. Montojo, D. Gerstenberger, and F. Xu, "Overview of 3GPP LTE-Advanced Carrier Aggregation for 4G wireless Communications, year=2012, month=February, volume=50, number=2, pages=122-130, doi=10.1109/MCOM.2012.6146491, issn=0163-6804,," *Communications Magazine, IEEE.*

[18] K. Pedersen, F. Frederiksen, C. Rosa, H. Nguyen, L. Garcia, and Y. Wang, "Carrier Aggregation for LTE-Advanced: Functionality and Performance Aspects," *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 89–95, June 2011.

[19] G. Yuan, X. Zhang, W. Wang, and Y. Yang, "Carrier Aggregation for LTE-Advanced Mobile Communication Systems," *Communications Magazine, IEEE*, vol. 48, no. 2, pp. 88–93, 2010.

[20] M. Iwamura, K. Etemad, M.-H. Fong, R. Nory, and R. Love, "Carrier Aggregation Framework in 3GPP LTE-Advanced [WiMAX/LTE Update]," *Communications Magazine, IEEE*, vol. 48, no. 8, pp. 60–67, August 2010.

[21] I. T. S. Sesia and M. Baker, *LTE the UMTS Long Term Evolution, from Theory to Practice*. Wiley Online Library, 2009.

[22] A. Larmo, M. Lindstrom, M. Meyer, G. Pelletier, J. Torsner, and H. Wiemann, "The LTE Link-Layer Design," *Communications Magazine, IEEE*, vol. 47, no. 4, pp. 52–59, April 2009.

[23] K. Pedersen, T. Kolding, F. Frederiksen, I. Kovacs, D. Laselva, and P. Mogensen, "An Overview of Downlink Radio Resource Management for UTRAN Long-Term Evolution," *Communications Magazine, IEEE*, vol. 47, no. 7, pp. 86–93, July 2009.

[24] A. Roessler and S. Merkel, "Carrier Aggregation: A Key Enabler for LTE-Advanced," *Microwave Journal*, vol. 55, no. 11, pp. 74–90, 2012.

[25] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2 (Release 10)," *TSG RAN*, vol. Release 10, TS 36.300 v10.3.0, 2011.

[26] Y. Wang, K. Pedersen, T. Sorensen, and P. Mogensen, "Carrier Load Balancing and Packet Scheduling for Multi-Carrier Systems," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1780–1789, May 2010.

[27] L. Zhang, F. Liu, L. Huang, and W. Wang, "Traffic Load Balance Methods in the LTE-Advanced System With Carrier Aggregation," in *Communications, Circuits and Systems (ICCCAS), 2010 International Conference on*, July 2010, pp. 63–67.

[28] L. Zhang, Y. Wang, L. Huang, H. Wang, and W. Wang, "QoS Performance Analysis on Carrier Aggregation Based LTE-A Systems," in *Wireless Mobile and Computing (CCWMC 2009), IET International Communication Conference on*, Dec 2009, pp. 253–256.

[29] Y. Wang, "System Level Analysis of LTE-Advanced: With Emphasis on Multi-Component Carrier Management," Ph.D. dissertation, Videnbasen for Aalborg UniversitetVBN, Aalborg UniversitetAalborg University, Det Teknisk-Naturvidenskabelige FakultetThe Faculty of Engineering and Science, Institut for Elektroniske SystemerDepartment of Electronic Systems, 2010.

[30] A. Khandekar, N. Bhushan, J. Tingfang, and V. Vanghi, "LTE-Advanced: Heterogeneous Networks," in *Wireless Conference (EW), 2010 European*, April 2010, pp. 978–982.

[31] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell Networks: a Survey," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59–67, September 2008.

[32] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. Thomas, J. Andrews, P. Xia, H. Jo, H. Dhillon, and T. Novlan, "Heterogeneous Cellular Networks: from Theory to Practice," *Communications Magazine, IEEE*, vol. 50, no. 6, pp. 54–64, June 2012.

[33] M.-S. Alouini and A. Goldsmith, "Area Spectral Efficiency of Cellular Mobile Radio Systems," *Vehicular Technology, IEEE Transactions on*, vol. 48, no. 4, pp. 1047–1066, Jul 1999.

[34] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A Survey on 3GPP Heterogeneous Networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 10–21, June 2011.

[35] H. Claussen, L. Ho, and L. Samuel, "Financial Analysis of a Pico-Cellular Home Network Deployment," in *Communications, 2007. ICC '07. IEEE International Conference on*, June 2007, pp. 5604–5609.

[36] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-Advanced: Next-Generation Wireless Broadband Technology [Invited Paper]," *Wireless Communications, IEEE*, vol. 17, no. 3, pp. 10–22, June 2010.

[37] M. Coupechoux, J.-M. Kelif, and P. Godlewski, "Network Controlled Joint Radio Resource Management for Heterogeneous Networks," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, May 2008, pp. 1771–1775.

[38] W. Song, H. Jiang, and W. Zhuang, "Performance Analysis of the WLAN-First Scheme in Cellular/WLAN Interworking," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 5, pp. 1932–1952, May 2007.

[39] A. Saleh, S. Redana, B. Raaf, and J. Hamalainen, "Comparison of Relay and Pico eNB Deployments in LTE-Advanced," in *Vehicular Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th*, Sept 2009, pp. 1–5.

[40] J. Andrews, "Seven Ways that HetNets are a Cellular Paradigm Shift," *Communications Magazine, IEEE*, vol. 51, no. 3, pp. 136–144, March 2013.

[41] B. Wang, H. Sun, R. Kapoor, S. Sambhwani, and M. Scipione, "Performance Analysis of HSPA Multi-Carrier Heterogeneous Networks," in *Communications (ICC), 2012 IEEE International Conference on*, June 2012, pp. 6035–6039.

[42] X. Lin, J. Andrews, and A. Ghosh, "Modeling, Analysis and Design for Carrier Aggregation in Heterogeneous Cellular Networks," *Communications, IEEE Transactions on*, vol. 61, no. 9, pp. 4002–4015, September 2013.

[43] 3GPP, "Technical Specication Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terres-

trial Radio Access Network (E-UTRAN); Overall Description; Stage 2(Release 11)," September 2012.

[44] T. Nihtila and V. Haikola, "HSDPA Performance with Dual Stream MIMO in a Combined Macro-Femto Cell Network," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, May 2010, pp. 1–5.

[45] H. R. Karimi, L. T. Ho, H. Claussen, and L. G. Samuel, "Evolution towards Dynamic Spectrum Sharing in Mobile Communications," in *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2006.

[46] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous Cellular Networks with Flexible Cell Association: A Comprehensive Downlink SINR Analysis," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 10, pp. 3484–3495, October 2012.

[47] K. Han, Y. Choi, D. Kim, M. Na, S. Choi, and K. Han, "Optimization of Femtocell Network Configuration under Interference Constraints," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on*, June 2009, pp. 1–7.

[48] J. Gora and T. Kolding, "Deployment Aspects of 3G Femtocells," in *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, Sept 2009, pp. 1507–1511.

[49] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference Coordination and Cancellation for 4G Networks," *Communications Magazine, IEEE*, vol. 47, no. 4, pp. 74–81, April 2009.

[50] F. Boccardi, J. Andrews, H. Elshaer, M. Dohler, S. Parkvall, P. Popovski, and S. Singh, "Why to Decouple the Uplink and Downlink in Cellular Networks and How To Do It," *arXiv preprint arXiv:1503.06746*, 2015.

[51] B. Soret, H. Wang, K. Pedersen, and C. Rosa, "Multicell Cooperation for LTE-Advanced Heterogeneous Network Scenarios," *Wireless Communications, IEEE*, vol. 20, no. 1, pp. 27–34, February 2013.

[52] A. Daeinabi, K. Sandrasegaran, and X. Zhu, "Performance Evaluation of Cell Selection Techniques for Picocells in LTE-Advanced Networks," in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2013 10th International Conference on*, May 2013, pp. 1–6.

[53] I. Guvenc, M.-R. Jeong, I. Demirdogen, B. Kecicioglu, and F. Watanabe, "Range Expansion and Inter-Cell Interference Coordination (ICIC) for Picocell Networks," in *Vehicular Technology Conference (VTC Fall), 2011 IEEE*, Sept 2011, pp. 1–6.

[54] 3GPP, "Study on Small Cell Enhancements for E-UTRA and E-UTRAN; Higher Layer Aspects," 3rd Generation Partnership Project (3GPP), TR 36.842, September 2014.

[55] Ericsson, "Further Discussions on UL/DL Split," 3GPP TSG-RAN, technical report R2-131678, May 2013.

[56] "3GPP TS 25.913: Group Radio Access Network - Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)."

[57] G. Piro, L. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating LTE Cellular Systems: An Open-Source Framework," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 2, pp. 498–513, 2011.

[58] D. Mcqueen, "The Momentum Behind LTE Adoption [sGPP LTE]," *Communications Magazine, IEEE*, vol. 47, no. 2, pp. 44–45, 2009.

[59] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 2, pp. 678–700, 2013.

[60] Z. Bakhti and S. Moghaddam, "Inter-Cell Interference Coordination with Adaptive Frequency-Reuse for VoIP and Data Traffic in Downlink of 3GPP-LTE," in *Application of Information and Communication Technologies (AICT), 2010 4th International Conference on*, 2010, pp. 1–6.

[61] S. Kumar, S. De, and H. Mohan, "Co-Operative Downlink Scheduling for Cell Edge and Handoff Users," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, 2012, pp. 2020–2024.

[62] R. Almatarneh, M. Ahmed, and O. Dobre, "Performance Analysis of Proportional Fair Scheduling in OFDMA Wireless Systems," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010, pp. 1–5.

[63] M. Ramesh Kumar, S. Bhashyam, and D. Jalihal, "Throughput Improvement for Cell-Edge Users Using Selective Cooperation in Cellular Networks," in *Wireless and Optical Communications Networks, 2008. WOCN '08. 5th IFIP International Conference on*, 2008.

[64] T. Innovations, "LTE in a Nutshell," *White paper*, 2010.

[65] J. Sanchez, D. Morales-Jimenez, G. Gomez, and J. Enbrambasaguas, "Physical Layer Performance of Long Term Evolution Cellular Technology," in *Mobile and Wireless Communications Summit, 2007. 16th IST*, July 2007, pp. 1–5.

[66] S.-w. Kim and K.-y. Kim, "Physical Layer Verification for 3GPP LTE (FDD)," in *Proceedings of the 11th international conference on Advanced Communication Technology-Volume 2*. IEEE Press, 2009, pp. 1095–1100.

[67] P. Frank, A. Muller, H. Droste, and J. Speidel, "Cooperative Interference-Aware Joint Scheduling for the 3GPP LTE Uplink," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, 2010, pp. 2216–2221.

[68] E. Tuomaala and H. Wang, "Effective SINR Approach of Link to System Mapping in OFDM/Multi-Carrier Mobile Network," in *Mobile Technology, Applications and Systems, 2005 2nd International Conference on*, Nov 2005, pp. 5 pp.–5.

[69] Z. Ma, Z. Cao, and W. Chen, "A Fair Opportunistic Spectrum Access (FOSA) Scheme in Distributed Cognitive Radio Networks," in *Communications, 2008. ICC '08. IEEE International Conference on*, 2008, pp. 4054–4058.

[70] J. Awad, M. Imran, and R. Tafazolli, "Improving Fairness by Cooperative Communications and Selection of Critical Users," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*, 2011, pp. 184–188.

[71] L. Akter and B. Natarajan, "Modeling fairness in resource allocation for secondary users in a competitive cognitive radio network," in *Wireless Telecommunications Symposium (WTS), 2010.*

[72] J. S. E. Dahlman, S. Parkvall and P. Beming, *3G Evolution HSPA and LTE for Mobile Broadband.* New York, USA: Academic Press, 2008.

[73] K. Younsun, J. Oh, H. Lee, K. Yongjun, J. Hyoungju, and K. Youngbum, "256 QAM Signal Transmission/Reception Method and Apparatus for Use in Mobile Communication System," Apr. 4 2014, uS Patent App. 14/245,423.

[74] D. H. Holma and D. A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*. Wiley Publishing, 2009.

[75] A. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Cooperative Routing in Static Wireless Networks," *Communications, IEEE Transactions on*, vol. 55, no. 11, pp. 2185–2192, 2007.

[76] 3GPP, "Dual-Cell High Speed Downlink Packet Access (HSDPA) Operation," 3GPP, technical report TR 25.825, June 2008.

[77] F. Haider, E. Hepsaydir, and N. Binucci, "Performance Analysis of LTE-Advanced Networks in Different Spectrum Bands," in *Wireless Advanced (WiAd), 2011*. IEEE, 2011, pp. 230–234.

[78] Y. Wang, K. Pedersen, T. Sorensen, and P. Mogensen, "Carrier Load Balancing and Packet Scheduling for Multi-Carrier Systems," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1780–1789, May 2010.

[79] H.-S. Liao, P.-Y. Chen, and W.-T. Chen, "An Efficient Downlink Radio Resource Allocation with Carrier Aggregation in LTE-Advanced Networks," pp. 1–1, 2014.

[80] Z. Huang, Y. Ji, and B. Zhao, "An Efficient Resource Allocation Algorithm with Carrier Aggregation in LTE Advanced Systems," in *Wireless Communications Signal Processing (WCSP), 2012 International Conference on*, Oct 2012, pp. 1–6.

[81] F. Wu, Y. Mao, X. Huang, and S. Leng, "A Joint Resource Allocation Scheme for OFDMA-Based Wireless Networks with Carrier Aggregation," in *WCNC, 2012 IEEE*, April 2012, pp. 1299–1304.

[82] H. Shajaiah, A. Abdel-Hadi, and C. Clancy, "Utility Proportional Fairness Resource Allocation with Carrier Aggregation in 4G-LTE," in *Military Communications Conference, MILCOM 2013 - 2013 IEEE*, Nov 2013, pp. 412–417.

[83] A. Abdelhadi and C. Clancy, "An Optimal Resource Allocation with Joint Carrier Aggregation in 4G-LTE," in *Computing, Networking and Communications (ICNC), 2015 International Conference on*, Feb 2015, pp. 138–142.

[84] A. Aijaz, X. Chu, and A. Aghvami, "Energy Efficient Design of SC-FDMA Based Uplink under QoS Constraints," *Wireless Communications Letters, IEEE*, vol. 3, no. 2, pp. 149–152, April 2014.

[85] M. Awad, V. Mahinthan, M. Mehrjoo, X. Shen, and J. W. Mark, "A Dual-Decomposition-Based Resource Allocation for OFDMA Networks With Imperfect CSI," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 5, pp. 2394–2403, Jun 2010.

[86] Y. Wang, K. Pedersen, T. Sorensen, and P. Mogensen, "Carrier Load Balancing and Packet Scheduling for Multi-Carrier Systems," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1780–1789, May 2010.

[87] S. Boyd and L. Vandenberghe, *Convex Optimization*, ser. Berichte über verteilte messysteme. Cambridge University Press, 2004. [Online]. Available: https://books.google.co.uk/books?id=mYm0bLd3fcoC

[88] D. Palomar and M. Chiang, "A Tutorial on Decomposition Methods for Network Utility Maximization," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 8, pp. 1439–1451, Aug 2006.

[89] W. Yu and R. Lui, "Dual Methods for Nonconvex Spectrum Optimization of Multi-Carrier Systems," *Communications, IEEE Transactions on*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[90] Y. Tachwali, B. Lo, I. Akyildiz, and R. Agusti, "Multiuser Resource Allocation Optimization Using Bandwidth-Power Product in Cognitive Radio Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 3, pp. 451–463, March 2013.

[91] J. Andrews, "Seven Ways that HetNets are a Cellular Paradigm Shift," *Communications Magazine, IEEE*, vol. 51, no. 3, pp. 136–144, March 2013.

[92] M. Arslan, K. Sundaresan, and S. Rangarajan, "Software-Defined Networking in Cellular Radio Access Networks: Potential and Challenges," *Communications Magazine, IEEE*, vol. 53, no. 1, pp. 150–156, January 2015.

[93] A. E. Lozano, R. W. H. Jr., and J. G. Andrews, "Fundamental Limits of Cooperation," *CoRR*, vol. abs/1204.0011, 2012. [Online]. Available: http://arxiv.org/abs/1204.0011

[94] J. Xu, Y. Guo, and R. Zhang, "CoMP Meets Energy Harvesting: A New Communication and Energy Cooperation Paradigm," in *Global Communications Conference (GLOBECOM), 2013 IEEE*, Dec 2013, pp. 2508–2513.

[95] P. Rost, A. Maeder, and X. Perez-Costa, "Asymmetric Uplink-Downlink Assignment for Energy-Efficient Mobile Communication Systems," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, May 2012, pp. 1–5.

[96] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, "LTE Release 12 and Beyond [Accepted From Open Call]," *Communications Magazine, IEEE*, vol. 51, no. 7, pp. 154–160, July 2013.

[97] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five Disruptive Technology Directions for 5G," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 74–80, February 2014.

[98] H. Elshaer, F. Boccardi, M. Dohler, and R. Irmer, "Downlink and Uplink Decoupling: A disruptive Architectural Design for 5G Networks," in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec 2014, pp. 1798–1803.

[99] K. Smiljkovikj, P. Popovski, and L. Gavrilovska, "Analysis of the Decoupled Access for Downlink and Uplink in Wireless Heterogeneous Networks," *Wireless Communications Letters, IEEE*, vol. PP, no. 99, pp. 1–1, 2015.

[100] A. de Domenico, V. Savin, and D. Ktenas, "A Backhaul-Aware Cell Selection Algorithm for Heterogeneous Cellular Networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, Sept 2013, pp. 1688–1693.

[101] F. Pantisano, M. Bennis, W. Saad, and M. Debbah, "Cache-Aware User Association in Backhaul-Constrained Small Cell Networks," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2014 12th International Symposium on*, May 2014, pp. 37–42.

[102] H. Elshaer, F. Boccardi, M. Dohler, and R. Irmer, "Load & Backhaul Aware Decoupled Downlink/Uplink Access in 5G Systems," *CoRR*, vol. abs/1410.6680, 2014. [Online]. Available: http://arxiv.org/abs/1410.6680

[103] K. Smiljkovikj, H. Elshaer, P. Popovski, F. Boccardi, M. Dohler, L. Gavrilovska, and R. Irmer, "Capacity Analysis of Decoupled Downlink and Uplink Access in 5G Heterogeneous Systems," *CoRR*, vol. abs/1410.7270, 2014. [Online]. Available: http://arxiv.org/abs/1410.7270

[104] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic Geometry for Modeling, Analysis, and Design of Multi-Tier and Cognitive Cellular Wireless Networks: A Survey," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 3, pp. 996–1019, Third 2013.

[105] S. Singh, X. Zhang, and J. G. Andrews, "Joint Rate and SINR Coverage Analysis for Decoupled Uplink-Downlink Biased Cell Associations in HetNets," *CoRR*, vol. abs/1412.1898, 2014. [Online]. Available: http://arxiv.org/abs/1412.1898